

Arg Ser Val Ala Lys Met Glu Ile Ala Arg Gln Gln Ser Cys Trp Leu
50 55 60

Val Cys Ile Tyr Cys Phe Arg Asn Pro Glu Ser Thr Leu Ala Pro Gly
65 70 75 80

Leu Pro Ala Cys Glu Ala Glu Leu Gly Leu Leu Arg Ala Gln Gly Leu
85 90 95

Pro His Pro Ala Ser Pro Ala Arg Leu Gly Asn Thr Gly Gly Ala Trp
100 105 110

Pro Arg Ser Lys Leu Gly Ser Gln Asn Thr Asn
115 120

<210> 385

<211> 26

<212> PRT

<213> Homo sapiens

<400> 385

Ser Ser Pro Ala Leu Ala Leu Thr Ser Pro Pro Lys Pro Leu Lys Gly
1 5 10 15

Glu Val Trp Leu Arg Trp Lys Leu Leu Gly
20 25

<210> 386

<211> 28

<212> PRT

<213> Homo sapiens

<400> 386

Glu His Lys Ala Tyr Pro Ile Leu Arg Leu Gln Pro Asp Leu Glu Thr
1 5 10 15

Gln Val Gly Pro Gly His Gly Val Asn Trp Asp Leu
20 25

<210> 387

<211> 28

<212> PRT

<213> Homo sapiens

<400> 387

Ala Leu Arg Cys Ser Leu Ser Cys Ser Leu Ile Pro Gly Leu Ser Pro
1 5 10 15

Asp Leu Ser Ser Glu Ala Pro Glu Gly Arg Ser Val
20 25

<210> 388

<211> 73

<212> PRT

<213> Homo sapiens

<400> 388

Leu Ala Pro Glu Cys Cys Cys Gly Ser Val Thr Tyr Pro Arg Ala Leu
1 5 10 15

Val Pro Arg Pro Cys Cys Pro Glu Pro Arg Ala Pro Leu Gln Leu Thr
20 25 30

Leu Gly Leu Phe Ser Ala Asn Pro Val Asn Ala Ser Pro Trp Gly Arg
35 40 45

Cys Arg Ser Arg Arg Gly Arg Gly Asn Leu Pro Leu Gly His Pro Val
50 55 60

Ser Thr Ala Phe Ser Ser Gly Asp Ser
65 70

<210> 389

<211> 102

<212> PRT

<213> Homo sapiens

<400> 389

Asn Thr Leu His Ser Lys Leu Val Pro Ser Val Tyr His Ser Thr Glu
1 5 10 15

Lys Ser Cys Leu Val Cys Phe Gly Met Cys Pro Ser Ile Tyr Lys Lys
20 25 30

Met Lys Ser Val Leu Leu Ile Gly Thr Arg Met Leu Leu Trp Leu Ser
35 40 45

His Ile Ser Gln Gly Pro Arg Pro Glu Ala Val Leu Pro Arg Ala Pro
50 55 60

Ser Pro Ser Ala Ala His Pro Trp Leu Val Phe Arg Lys Pro Gly Lys
65 70 75 80

Arg Lys Pro Leu Gly Gln Met Gln Lys Gln Lys Arg Glu Gly Lys Pro
85 90 95

Ala Ser Gly Ser Pro Cys
100

<210> 390

<211> 25

<212> PRT

<213> Homo sapiens

<400> 390

Tyr Pro Arg Ala Leu Val Pro Arg Pro Cys Cys Pro Glu Pro Arg Ala
1 5 10 15

Pro Leu Gln Leu Thr Leu Gly Leu Phe

20

25

<210> 391
<211> 27
<212> PRT
<213> Homo sapiens

<400> 391
Val Leu Leu Ile Gly Thr Arg Met Leu Leu Trp Leu Ser His Ile Ser
1 5 10 15
Gln Gly Pro Arg Pro Glu Ala Val Leu Pro Arg
20 25

<210> 392
<211> 61
<212> PRT
<213> Homo sapiens

<400> 392
Trp Ile Ile Val Met Phe Gly Lys Val Leu Lys Ile Lys Asp Phe Met
1 5 10 15
Ser Thr Tyr Ser His Thr Tyr Thr His Thr His Met His Ala His Thr
20 25 30
His Thr His Thr Leu Thr Leu Ser Leu Leu Gln Asn Val Leu Thr Leu
35 40 45
Val Ala Ile Ser Asp Ser Asp Lys Ala Leu Leu Ile Phe
50 55 60

<210> 393
<211> 69
<212> PRT
<213> Homo sapiens

<400> 393
Met Thr Leu Leu Ile Ala Glu Lys Thr Trp Arg Arg Pro Trp Pro Cys
1 5 10 15
Gln Trp Gly Tyr Leu Gly Ala Glu Gly Asp Arg His Leu Glu Gly Arg
20 25 30
Ser Leu Ser Leu Arg His Leu Gln Gly Ala Glu Thr Pro Val Leu Asn
35 40 45
Pro Asp Leu Gln Leu Pro Ser His Ile Gly Lys Gln Ala Trp Ser His
50 55 60
Ala Leu Gly Ser Leu
65

<210> 394

<211> 27
<212> PRT
<213> Homo sapiens

<400> 394

Met Ser Thr Tyr Ser His Thr Tyr Thr His Thr His Met His Ala His
1 5 10 15

Thr His Thr His Thr Leu Thr Leu Ser Leu Leu
20 25

<210> 395

<211> 23

<212> PRT

<213> Homo sapiens

<400> 395

Gly Ala Glu Gly Asp Arg His Leu Glu Gly Arg Ser Leu Ser Leu Arg
1 5 10 15

His Leu Gln Gly Ala Glu Thr
20

<210> 396

<211> 133

<212> PRT

<213> Homo sapiens

<400> 396

Val Val Glu Pro Gly Leu Lys Ala Ser Leu Gly Ala Met Ser Thr Leu
1 5 10 15

Phe Pro Ser Leu Phe Pro Arg Val Thr Glu Thr Leu Trp Phe Asn Leu
20 25 30

Asp Arg Pro Cys Val Glu Glu Thr Glu Leu Gln Gln Gln Glu Gln Gln
35 40 45

His Gln Ala Trp Leu Gln Ser Ile Ala Glu Lys Asp Asn Asn Leu Val
50 55 60

Pro Ile Gly Lys Pro Ala Ser Glu His Tyr Asp Asp Glu Glu Glu Glu
65 70 75 80

Asp Asp Glu Asp Asp Glu Asp Ser Glu Glu Asp Ser Glu Asp Asp Glu
85 90 95

Asp Met Gln Asp Met Asp Glu Met Asn Asp Tyr Asn Glu Ser Pro Asp
100 105 110

Asp Gly Glu Val Asn Glu Val Asp Met Glu Gly Asn Glu Gln Asp Gln
115 120 125

Asp Gln Trp Met Ile
130

<210> 397
<211> 23
<212> PRT
<213> Homo sapiens

<400> 397
Leu Phe Pro Arg Val Thr Glu Thr Leu Trp Phe Asn Leu Asp Arg Pro
1 5 10 15

Cys Val Glu Glu Thr Glu Leu
20

<210> 398
<211> 23
<212> PRT
<213> Homo sapiens

<400> 398
Tyr Asn Glu Ser Pro Asp Asp Gly Glu Val Asn Glu Val Asp Met Glu
1 5 10 15

Gly Asn Glu Gln Asp Gln Asp
20

<210> 399
<211> 101
<212> PRT
<213> Homo sapiens

<400> 399
Met Gly Phe Asp Ile His Gly Val Leu Gly Glu Ala Val Ala Glu Pro
1 5 10 15

Arg Glu Lys Lys Gln Glu Arg Ala Lys Trp Ala Pro His Asp Tyr Asp
20 25 30

Asp Pro Ser Leu Ser Leu Gln Asp Leu Leu Ile Ser Trp Met Ile Ser
35 40 45

Thr Trp Leu Ile Pro Met Trp Lys Cys Gln Ala Thr Ile Trp Phe Ser
50 55 60

Leu Ile Gln Arg Leu Leu Asn Ala Tyr Cys Met Pro Gly Asn Phe Arg
65 70 75 80

His Trp Glu Ile Ala Ala Asn Thr Thr Asn Lys Thr Pro Gly Leu Met
85 90 95

Asp Phe Lys Phe Leu
100

<210> 400
<211> 27
<212> PRT

<213> Homo sapiens

<400> 400

Glu Pro Arg Glu Lys Lys Gln Glu Arg Ala Lys Trp Ala Pro His Asp
1 5 10 15

Tyr Asp Asp Pro Ser Leu Ser Leu Gln Asp Leu
20 25

<210> 401

<211> 24

<212> PRT

<213> Homo sapiens

<400> 401

Met Pro Gly Asn Phe Arg His Trp Glu Ile Ala Ala Asn Thr Thr Asn
1 5 10 15

Lys Thr Pro Gly Leu Met Asp Phe
20

<210> 402

<211> 100

<212> PRT

<213> Homo sapiens

<400> 402

Gln Ser Val Pro Ser Pro Pro Leu Ala Pro Pro Leu Pro Pro Ser Leu
1 5 10 15

Pro Ser Phe Leu Phe Thr Glu Thr Arg Ser His Tyr Val Ala Arg Leu
20 25 30

Val Ser Asn Ser Trp Ala Gln Met Ile Leu Leu Pro Trp Pro Leu Lys
35 40 45

Val Leu Gly Leu Asp Val Ser His Cys Ala Trp Pro Lys Ser Val Phe
50 55 60

Leu Gln Ala Met Glu Glu Ile Ala Asp Phe Cys Leu Phe Ser Val Lys
65 70 75 80

Tyr Gln Val Ser Ser Met Thr Cys Phe Asp Arg Thr Ser Tyr Met Lys
85 90 95

Asn Thr Tyr Leu
100

<210> 403

<211> 27

<212> PRT

<213> Homo sapiens

<400> 403

Leu Phe Thr Glu Thr Arg Ser His Tyr Val Ala Arg Leu Val Ser Asn

1

5

10

15

Ser Trp Ala Gln Met Ile Leu Leu Pro Trp Pro
 20 25

<210> 404

<211> 159

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (124)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (142)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 404

Ser Gln Ile Lys Ser Glu Lys Lys His Ile Gly Lys Ala Tyr Thr Cys
 1 5 10 15

Thr Gln Thr Gln Ser Thr Gly Met Gln Ser Thr Leu Thr Ile Val Ala
 20 25 30

Lys Lys Lys Ser Arg Asn His Thr Glu Ser Tyr Thr Arg Lys Lys Gln
 35 40 45

Glu Asn Gln Ile Val Leu Ile Pro Trp His Gln Lys Lys His Pro Glu
 50 55 60

Gly Thr His Thr Cys Ser His Ser Leu Arg Arg Asp Thr Asn Thr Ala
 65 70 75 80

Ala Asp Thr Gln Arg Lys Ile Arg Ala His Arg Tyr Thr Tyr Arg Arg
 85 90 95

Asp Lys Tyr Ser Asp Thr Leu Val Thr His Asp His Tyr Lys Gly Asp
 100 105 110

Lys His Pro Ser Asn Thr His Thr Gln Pro Arg Xaa Glu Phe Leu Gln
 115 120 125

Pro Gly Gly Ser Thr Asn Ser Arg Ala Ala Ala Pro Arg Xaa Ser Ser
 130 135 140

Ser Phe Cys Pro Phe Ser Glu Gly Tyr Ser Ser Trp Gly Tyr His
 145 150 155

<210> 405

<211> 26

<212> PRT

<213> Homo sapiens

<400> 405

Gly Met Gln Ser Thr Leu Thr Ile Val Ala Lys Lys Lys Ser Arg Asn
1 5 10 15

His Thr Glu Ser Tyr Thr Arg Lys Lys Gln
20 25

<210> 406

<211> 24

<212> PRT

<213> Homo sapiens

<400> 406

Lys Lys His Pro Glu Gly Thr His Thr Cys Ser His Ser Leu Arg Arg
1 5 10 15

Asp Thr Asn Thr Ala Ala Asp Thr
20

<210> 407

<211> 24

<212> PRT

<213> Homo sapiens

<400> 407

Arg Arg Asp Lys Tyr Ser Asp Thr Leu Val Thr His Asp His Tyr Lys
1 5 10 15

Gly Asp Lys His Pro Ser Asn Thr
20

<210> 408

<211> 91

<212> PRT

<213> Homo sapiens

<400> 408

Lys His Leu Pro Leu Lys Ala Pro Ile Asp Leu Asp Asn Lys Asn Ser
1 5 10 15

Cys Met Phe Cys Ser Arg Asp Ile Phe Cys Arg Phe His His Ser Thr
20 25 30

Ala Trp Leu Phe Leu Gly Arg Ile Thr Asp Arg Ile Leu Gly Leu His
35 40 45

His Tyr Leu Ile Arg Tyr Gln Phe Glu Ile Glu Asn Leu Cys Leu Met
50 55 60

Lys Ile Val Ile Pro Val Val Ser Met Lys Thr Asn Cys Gln Phe Asp
65 70 75 80

Phe Leu Gly Gln Leu Lys Gln Asn Leu Tyr His
85 90

<210> 409
<211> 28
<212> PRT
<213> Homo sapiens

<400> 409
Ile Glu Asn Leu Cys Leu Met Lys Ile Val Ile Pro Val Val Ser Met
1 5 10 15
Lys Thr Asn Cys Gln Phe Asp Phe Leu Gly Gln Leu
20 25

<210> 410
<211> 21
<212> PRT
<213> Homo sapiens

<400> 410
Ala Pro Ile Asp Leu Asp Asn Lys Asn Ser Cys Met Phe Cys Ser Arg
1 5 10 15
Asp Ile Phe Cys Arg
20

<210> 411
<211> 53
<212> PRT
<213> Homo sapiens

<400> 411
Gly Thr Ser Val Asn Glu Ser Val Ser Asn Ala Thr Ala Ile Asp Ser
1 5 10 15
Gln Ile Ala Arg Ser Leu His Ile Pro Leu Thr Gln Asp Ile Ala Gly
20 25 30
Asp Pro Ser Tyr Glu Ile Ser Lys Gln Arg Leu Ser Ile Val Ile Gly
35 40 45
Val Val Ala Gly Ile
50

<210> 412
<211> 220
<212> PRT
<213> Homo sapiens

<400> 412
Pro Lys Ile Lys Met Ala Met Lys Pro Ala Lys Lys Ile Thr Lys Thr
1 5 10 15
Phe Leu His Pro Asn Ser Met Thr Asn Leu Lys Ser Leu Lys Arg Thr
20 25 30

Arg Lys Thr Lys Asn Leu Ser Ser Leu Ser Thr Ala Ala Leu Ser Leu
 35 40 45
 Trp Arg Leu Leu Ser Gln Met Asp Arg Gly Met Ile Val Ser Met Arg
 50 55 60
 Ser Cys Gln Thr Ala Gln Ala Trp Gly Asp Thr Gly Pro Leu Met Val
 65 70 75 80
 Gly Pro Ala Val Leu Thr Trp Gln Gly Ile Thr Asn Leu Val Pro His
 85 90 95
 Cys Leu Leu Phe Ser Phe Ile Pro Ser His Gln Leu Gln Glu Lys Asn
 100 105 110
 Thr Arg Pro Tyr Lys Ile Tyr His Gln Pro Thr His Leu Trp Glu Gln
 115 120 125
 Glu Thr Thr Phe Gln Leu Asp Gln Ile Thr Ala Leu Ser Thr Ala Val
 130 135 140
 Lys Pro Ile Thr Ser Thr Ala Asn Arg Cys Val Tyr Ile His Thr Leu
 145 150 155 160
 Leu Cys Leu Ala Glu Phe His Ser Asn Met Met Leu His Tyr Ala Pro
 165 170 175
 Tyr Cys Asp Asp Leu Ser Thr Pro Lys Pro Ala Gly Ala Cys Pro Trp
 180 185 190
 Pro Trp Gly Val Ser Gln Ser Leu Leu Val Pro Leu Val Val His Phe
 195 200 205
 Ile Phe Glu Ser Phe Ser Phe Ser Tyr Thr Glu Lys
 210 215 220

<210> 413

<211> 55

<212> PRT

<213> Homo sapiens

<400> 413

Cys Ser Ile Met His His Thr Val Met Thr Phe Leu Leu Arg Asn Leu
 1 5 10 15
 Leu Glu Pro Ala Leu Gly Arg Gly Val Ser Ala Asn His Cys Leu Phe
 20 25 30
 His Leu Leu Tyr Ile Leu Phe Leu Ser Leu Phe Leu Ser His Ile Gln
 35 40 45
 Lys Asn Ser Met Lys Ile Lys
 50 55

<210> 414

<211> 29

<212> PRT
<213> Homo sapiens

<400> 414
Thr Ala Ile Asp Ser Gln Ile Ala Arg Ser Leu His Ile Pro Leu Thr
1 5 10 15
Gln Asp Ile Ala Gly Asp Pro Ser Tyr Glu Ile Ser Lys
20 25

<210> 415
<211> 21
<212> PRT
<213> Homo sapiens

<400> 415
Tyr Cys Arg Ser Lys Asn Lys Asn Gly Tyr Glu Ala Gly Lys Lys Asp
1 5 10 15
His Glu Asp Phe Phe
20

<210> 416
<211> 21
<212> PRT
<213> Homo sapiens

<400> 416
Gly Pro Gly Ser Pro Asp Leu Ala Arg His Tyr Lys Ser Ser Ser Pro
1 5 10 15
Leu Pro Thr Val Gln
20

<210> 417
<211> 25
<212> PRT
<213> Homo sapiens

<400> 417
Leu Pro Pro Ala Asn Thr Phe Val Gly Ala Gly Asp Asn Ile Ser Ile
1 5 10 15
Gly Ser Asp His Cys Ser Glu Tyr Ser
20 25

<210> 418
<211> 119
<212> PRT
<213> Homo sapiens

<400> 418
Gly Thr Ser Asn Ala Ser Val Ser Pro Thr Ile Cys Ile Cys Met Cys
1 5 10 15

Gly Tyr Val His Ile Trp Phe Phe Ile Cys Leu Cys Val Tyr Leu Lys
20 25 30

Val Leu Gln Gly Ser Ala Cys Pro Trp Ile Ala Ala Ala Val Val Met
35 40 45

Arg Arg Met Arg Lys Val Gln Glu Lys Gly Glu Val Phe Arg Asn Met
50 55 60

Ala Ala Thr Trp Ala Leu Arg Ser Gly Ile Gln Ser Leu Asn Ser Leu
65 70 75 80

Val Ser Ser Ala Phe Phe Thr Ile Phe Met Thr Leu Gly Ser Ser Trp
85 90 95

Asn Leu Ile Val Ser Leu Ser Ser Leu Val Asn Trp Thr Gly Leu Phe
100 105 110

Ser Phe Tyr Phe Ser Arg Asn
115

<210> 419

<211> 28

<212> PRT

<213> Homo sapiens

<400> 419

Cys Leu Cys Val Tyr Leu Lys Val Leu Gln Gly Ser Ala Cys Pro Trp
1 5 10 15

Ile Ala Ala Ala Val Val Met Arg Arg Met Arg Lys
20 25

<210> 420

<211> 26

<212> PRT

<213> Homo sapiens

<400> 420

Thr Ile Phe Met Thr Leu Gly Ser Ser Trp Asn Leu Ile Val Ser Leu
1 5 10 15

Ser Ser Leu Val Asn Trp Thr Gly Leu Phe
20 25

<210> 421

<211> 58

<212> PRT

<213> Homo sapiens

<400> 421

Gln Pro Asp Ile Pro Val Leu Pro Val Gly Phe Ser Gln Asn Cys Ser
1 5 10 15

Phe Lys Val Ser Gly Cys Trp Lys Gly Gly Leu Ile Ala Glu Lys Val
 20 25 30

Gly Thr Leu Gly Thr Pro Lys Gly Arg Arg Ala Trp Pro Glu Thr Glu
 35 40 45

Phe Phe Arg Phe Leu Glu Pro Gly Leu Pro
 50 55

<210> 422

<211> 131

<212> PRT

<213> Homo sapiens

<400> 422

Arg Gly Phe Arg Met Ala Gln Pro Leu Val Asn Thr Phe Gln Val Ala
 1 5 10 15

Val Pro Val Glu Asp Leu Ala Pro Gln Gln Asn Pro Ser Arg Phe Pro
 20 25 30

Ala Asp Pro Ala Leu Leu Ser Phe Leu Thr Gly Ser Ile Leu Ala Pro
 35 40 45

Gly Lys Val Ile Trp Val Asn Val Ser Phe Thr Ala Ile Ile Trp Pro
 50 55 60

Thr Trp Asp Ser Met Ala Ile Gly Glu Leu Thr Ile Ala Ser His Ala
 65 70 75 80

Ser Met Thr Leu His Ile Gly Arg Pro Gly Ser Arg Lys Arg Lys Asn
 85 90 95

Ser Val Ser Gly His Ala Arg Leu Pro Phe Gly Val Pro Ser Val Pro
 100 105 110

Thr Phe Ser Ala Ile Ser Pro Pro Phe Gln Gln Pro Glu Thr Leu Lys
 115 120 125

Glu Gln Phe
 130

<210> 423

<211> 24

<212> PRT

<213> Homo sapiens

<400> 423

Glu Asp Leu Ala Pro Gln Gln Asn Pro Ser Arg Phe Pro Ala Asp Pro
 1 5 10 15

Ala Leu Leu Ser Phe Leu Thr Gly
 20

<210> 424

<211> 29
<212> PRT
<213> Homo sapiens

<400> 424

Thr Trp Asp Ser Met Ala Ile Gly Glu Leu Thr Ile Ala Ser His Ala
1 5 10 15

Ser Met Thr Leu His Ile Gly Arg Pro Gly Ser Arg Lys
20 25

<210> 425
<211> 71
<212> PRT
<213> Homo sapiens

<400> 425

Val Ser Pro Gln Leu Met Gly Ile Lys Arg Glu Pro Ser Ala Ala Gln
1 5 10 15

Leu Ser Val Gly Glu Glu His Thr Leu Asp Arg Glu Gly Arg Glu Leu
20 25 30

Val Asp Leu Pro Gly Gln Pro Ser Gln Lys Ile Lys Ile Lys Asn Lys
35 40 45

Ser Ser Leu His Pro Gly Leu Ile Ile Pro Pro Ala His Tyr Lys Thr
50 55 60

Ala Thr Thr Thr Asn Leu Phe
65 70

<210> 426
<211> 21
<212> PRT
<213> Homo sapiens

<400> 426

Pro Ser Ala Ala Gln Leu Ser Val Gly Glu Glu His Thr Leu Asp Arg
1 5 10 15

Glu Gly Arg Glu Leu
20

<210> 427
<211> 23
<212> PRT
<213> Homo sapiens

<400> 427

Asn Cys Asp His Asp Phe Ile Gln Pro Leu His Thr Pro Met Ser Ala
1 5 10 15

Leu Phe Gln Ser Glu Phe Ser
20

<210> 428
<211> 107
<212> PRT
<213> Homo sapiens

<400> 428
Ser Ile Leu Asn Met Gly Leu Phe Thr Glu Gln Arg Pro Trp Pro Ala
1 5 10 15
Ala Ala Arg Cys Ala Arg Gln Ser Thr Val Ala Gly Ala Ile Arg Arg
20 25 30
Ala Arg Gly Thr Val Thr Met Trp Gln Val Ala Gly Ala Ala Trp Ala
35 40 45
Ser Pro Asp Arg Arg Ala Lys Val His Pro Cys Arg His Ala Ala Pro
50 55 60
Cys Leu Pro Ser Pro Cys Arg Arg Gly Leu Gln Met Ser Gly Pro Leu
65 70 75 80
Gln Ala Thr Arg Gly Arg Val Thr Leu Arg Ser His Gln Val Gly Cys
85 90 95
Lys Arg Ala Thr Gly Ser Ile Glu Asn Ser Leu
100 105

<210> 429
<211> 114
<212> PRT
<213> Homo sapiens

<400> 429
Gln Lys Ser Lys Gly Ser Pro Leu Gln Thr Cys Cys Ser Leu Pro Thr
1 5 10 15
Leu Pro Met Gln Glu Arg Pro Ala Asp Glu Trp Ser Thr Pro Gly Asp
20 25 30
Gln Gly Lys Ser Tyr Ile Lys Lys Pro Pro Gly Gly Leu Gln Lys Gly
35 40 45
His Arg Leu His Arg Lys Leu Thr Leu Lys Gln Gly Arg His Arg Gly
50 55 60
Val Glu Gly Leu Asn Glu Ile Met Val Thr Val Leu Lys Glu Glu Phe
65 70 75 80
Pro Val Ser Lys Pro Gly Leu Asn Val Leu Pro Thr Phe His Arg His
85 90 95
His Glu Cys Tyr Gln His Gly Met Asn Leu Thr Ala Arg Ile Ser Val
100 105 110
Val Ser

<210> 430
<211> 25
<212> PRT
<213> Homo sapiens

<400> 430
Ala Arg Gln Ser Thr Val Ala Gly Ala Ile Arg Arg Ala Arg Gly Thr
1 5 10 15
Val Thr Met Trp Gln Val Ala Gly Ala
20 25

<210> 431
<211> 25
<212> PRT
<213> Homo sapiens

<400> 431
Pro Cys Arg Arg Gly Leu Gln Met Ser Gly Pro Leu Gln Ala Thr Arg
1 5 10 15
Gly Arg Val Thr Leu Arg Ser His Gln
20 25

<210> 432
<211> 26
<212> PRT
<213> Homo sapiens

<400> 432
Leu Pro Met Gln Glu Arg Pro Ala Asp Glu Trp Ser Thr Pro Gly Asp
1 5 10 15
Gln Gly Lys Ser Tyr Ile Lys Lys Pro Pro
20 25

<210> 433
<211> 23
<212> PRT
<213> Homo sapiens

<400> 433
Asn Val Leu Pro Thr Phe His Arg His His Glu Cys Tyr Gln His Gly
1 5 10 15
Met Asn Leu Thr Ala Arg Ile
20

<210> 434
<211> 40
<212> PRT

<213> Homo sapiens

<400> 434

Ile Asn Val Leu Tyr Cys Ser Arg Asp Ser Leu Met Gly Arg Thr Ile
1 5 10 15

Met Glu Ser Ser Asp Tyr Ile Lys Lys Gly Ala Asn Val Ser Pro Val
20 25 30

Leu Gly Val Arg Gln Gln Ala Val
35 40

<210> 435

<211> 28

<212> PRT

<213> Homo sapiens

<400> 435

Ser Leu Leu Met Tyr Phe Val Phe Lys Ile Phe Phe Gln Ser Leu Cys
1 5 10 15

Val Leu Gly Tyr Cys Ile Leu Pro Leu Thr Val Ala
20 25

<210> 436

<211> 50

<212> PRT

<213> Homo sapiens

<400> 436

Arg Leu Trp Met Thr Lys Ala His Pro Ala Leu Arg His Leu Leu Leu
1 5 10 15

Leu Phe Thr Leu Ala Leu Thr Leu Leu Ala Gln Gly Cys Cys Ala Val
20 25 30

Ala Pro Ser Gly Cys Ala Asp Leu Ala Gly Phe Cys Ser Leu Gly His
35 40 45

Ser Cys
50

<210> 437

<211> 48

<212> PRT

<213> Homo sapiens

<400> 437

Arg Thr Cys Thr Pro Trp Met Gly Phe Trp Cys Leu Val Cys Ser Leu
1 5 10 15

Phe Ala Pro Val Pro Thr Ser Arg Lys Tyr Leu Val Ser Lys Pro Gly
20 25 30

Cys Tyr Gln Arg Arg Arg Val Phe Gly Val Cys Phe Thr Lys Pro Leu

35

40

<210> 438
<211> 8
<212> PRT
<213> Homo sapiens

<400> 438
Trp Leu Leu Ser Glu Lys Lys Gly
1 5

<210> 439
<211> 10
<212> PRT
<213> Homo sapiens

<400> 439
Gly Val Phe Tyr Lys Ala Ala Val Ile Gly
1 5 10

<210> 440
<211> 45
<212> PRT
<213> Homo sapiens

<400> 440
Cys Lys Thr Ser Pro Leu Pro Lys Glu Gly Gln Ser Ala Val Ser Val
1 5 10 15

Pro Val Ser Ser His Phe Leu Ala His Ser Ala Pro Leu Ser Gly Gly
20 25 30

His Ala His Val Phe Ala Arg Asp Gly Ala Thr Gly Leu
35 40 45

<210> 441
<211> 140
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (54)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 441
Leu Gly Arg Gly Ser Gly Glu Arg Lys Thr Pro Val Ser Cys Phe Ala
1 5 10 15

Gln Ile Ser Lys Ser Arg Gly Gly Arg Ser Lys Ser Leu Thr His Leu
20 25 30

Cys Thr His Thr His Thr Gln Val Thr Glu Leu Asp Val Arg Met Ser
35 40 45

His Gly Cys Leu Arg Xaa Gln His Ala Gly Arg Leu Ala Pro Pro Pro
50 55 60

Pro Leu Arg Phe Cys Leu Thr Ala Cys Trp Gly Arg Arg Gly Glu Ala
65 70 75 80

Glu Thr Val Trp Lys Asp Pro Ala Ser Ser Gln His Pro Pro Pro Ser
85 90 95

Glu Lys Pro His Arg Gln Asp Arg His Pro Glu Arg Trp His Gln Pro
100 105 110

Gly Gly Pro Ile Pro Gly Lys His Met Arg Val Ser Pro Gly Gln Arg
115 120 125

Gly Arg Val Cys Gln Glu Met Gly Arg Asn Arg Asn
130 135 140

<210> 442

<211> 102

<212> PRT

<213> Homo sapiens

<400> 442

Phe Cys Leu Arg Asp Phe Lys Ile Trp Arg Gly Arg Leu Glu Ala Gly
1 5 10 15

Arg Thr Glu Gly Arg Leu Ala Gly Glu Arg Phe Gly Gly Glu Glu Asp
20 25 30

Pro Ser Phe Leu Phe Cys Ser Asp Phe Lys Val Glu Gly Trp Ala Phe
35 40 45

Glu Ile Ser His Ser Leu Val His Thr His Thr His Thr Gly His Gly
50 55 60

Ala Gly Arg Ala Asp Val Thr Arg Val Pro Ala Gly Thr Ala Arg Trp
65 70 75 80

Glu Ala Gly Ser Pro Thr Pro Ser Pro Val Leu Phe Asp Ser Leu Leu
85 90 95

Gly Ala Ala Gly Arg Gly
100

<210> 443

<211> 28

<212> PRT

<213> Homo sapiens

<400> 443

Ala Gln Ile Ser Lys Ser Arg Gly Gly Arg Ser Lys Ser Leu Thr His

1

10

15

Leu Cys Thr His Thr His Thr Gln Val Thr Glu Leu
20 25

<210> 444

<211> 26

<212> PRT

<213> Homo sapiens

<400> 444

Glu Lys Pro His Arg Gln Asp Arg His Pro Glu Arg Trp His Gln Pro
1 5 10 15

Gly Gly Pro Ile Pro Gly Lys His Met Arg
20 25

<210> 445

<211> 26

<212> PRT

<213> Homo sapiens

<400> 445

Gly Arg Leu Glu Ala Gly Arg Thr Glu Gly Arg Leu Ala Gly Glu Arg
1 5 10 15

Phe Gly Gly Glu Glu Asp Pro Ser Phe Leu
20 25

<210> 446

<211> 23

<212> PRT

<213> Homo sapiens

<400> 446

Val Thr Arg Val Pro Ala Gly Thr Ala Arg Trp Glu Ala Gly Ser Pro
1 5 10 15

Thr Pro Ser Pro Val Leu Phe
20

<210> 447

<211> 31

<212> PRT

<213> Homo sapiens

<400> 447

Asp Glu Gly Val Gln Gly Glu Arg Leu Phe Arg Ile Leu Arg Ile Asn
1 5 10 15

Gly Glu Lys Pro Tyr Asn Phe Val Asp Tyr Phe His Cys Glu Tyr
20 25 30

<210> 448
 <211> 111
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (59)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (62)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (65)
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>
 <221> SITE
 <222> (66)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 448
 Lys Val Val Arg Ile Asp Asn Gly Ile Leu Cys Ser His Lys Lys Thr
 1 5 10 15
 Glu Ile Met Ser Leu Gln Gln His Gly Trp Ile Trp Arg Pro Tyr Leu
 20 25 30
 Lys Gln Thr Asn Thr Gly Thr Glu Asn Gln Ile Pro His Thr Leu Thr
 35 40 45
 Tyr Lys Trp Glu Leu Asn Phe Glu Tyr Ile Xaa Thr Gln Xaa Arg Gly
 50 55 60
 Xaa Xaa Asp Ser Glu Ala Tyr Leu Lys Val Glu Gly Gly Arg Arg Glu
 65 70 75 80
 Gly Ile Gln Lys Leu Pro Ile Arg Tyr Tyr Val Tyr Tyr Leu Gly Asp
 85 90 95
 Lys Ile Ile Cys Thr Ser Ser Ser Cys Ser Met His Leu Leu Met
 100 105 110

<210> 449
 <211> 21
 <212> PRT
 <213> Homo sapiens

<400> 449
 His Lys Asp Thr Cys Met Ser Met Phe Thr Ala Ala Leu Phe Thr Ile
 1 5 10 15
 Ala Lys Thr Trp Asn

20

<210> 450
<211> 14
<212> PRT
<213> Homo sapiens

<400> 450
Met Pro Ile Asn Asp Arg Leu Asp Phe Lys Arg Trp Tyr Val
1 5 10

<210> 451
<211> 47
<212> PRT
<213> Homo sapiens

<400> 451
Thr Met Glu Ser Tyr Val Ala Ile Lys Arg Gln Arg Ser Cys Pro Cys
1 5 10 15

Ser Asn Met Val Gly Ser Gly Gly His Ile Leu Ser Lys Leu Thr Gln
20 25 30

Glu Gln Lys Thr Lys Tyr His Ile Leu Ser Leu Ile Ser Gly Ser
35 40 45

<210> 452
<211> 25
<212> PRT
<213> Homo sapiens

<400> 452
Glu Ile Met Ser Leu Gln Gln His Gly Trp Ile Trp Arg Pro Tyr Leu
1 5 10 15

Lys Gln Thr Asn Thr Gly Thr Glu Asn
20 25

<210> 453
<211> 24
<212> PRT
<213> Homo sapiens

<400> 453
Arg Arg Glu Gly Ile Gln Lys Leu Pro Ile Arg Tyr Tyr Val Tyr Tyr
1 5 10 15

Leu Gly Asp Lys Ile Ile Cys Thr
20

<210> 454
<211> 57
<212> PRT

<213> Homo sapiens

<400> 454

Leu His Gly Glu Gln Val Pro Ile Tyr Ile Phe Leu Leu Met Gln Pro
1 5 10 15

Leu Asn Phe Glu Cys Ile Ser Phe Leu Asn Cys Ile Glu Gln Tyr Ser
20 25 30

Val Gly Val Ile His Asn Ser Val Thr Ile Tyr Ala Cys Asp Arg Glu
35 40 45

Glu Asn Cys Met Asp Ile Arg Tyr Leu
50 55

<210> 455

<211> 12

<212> PRT

<213> Homo sapiens

<400> 455

Gly Thr Ser Trp Ala Ser Arg Phe Phe Thr Cys His
1 5 10

<210> 456

<211> 52

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (5)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (7)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (15)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (18)

<223> Xaa equals any of the naturally occurring L-amino acids

<220>

<221> SITE

<222> (37)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 456

Gly Pro Pro Arg Xaa Phe Xaa Pro Lys Lys Ala Ile Leu Gly Xaa Pro

1

5

10

15

Pro Xaa Gly Arg Val Pro Pro Phe Arg Tyr Arg Ser Arg Asn Ser Arg
20 25 30

Gly Arg Pro His Xaa Ser Ala Pro Arg Val Arg Phe Cys Leu Glu Asn
35 40 45

Ser Trp Leu Arg
50

<210> 457

<211> 72

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (56)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 457

Pro Leu Asn Thr Met Met Cys Met Met Cys Lys Met Lys Val Ser Pro
1 5 10 15

Lys Ile Phe Ser Lys Leu Lys Arg Lys Tyr Leu Asn Ser Asn Thr Leu
20 25 30

Thr Lys Leu Glu Met Gln Thr Val His Leu Glu Ser Ser Leu Ala Ser
35 40 45

Cys Ser Pro Asn Lys Ser Gly Xaa Val Gly Arg Thr Arg Gly Val Asp
50 55 60

Pro Gly Asn Ser Gly Thr Gly Thr
65 70

<210> 458

<211> 69

<212> PRT

<213> Homo sapiens

<400> 458

Gly Thr Val Thr Gln Lys Arg Lys Cys Val Phe Gly Lys Tyr Leu Leu
1 5 10 15

Ser Thr Cys Ser Leu Met Phe Ser Ser Met His Gly Ala Cys Ser Trp
20 25 30

Lys Ala Lys Gln Thr Ser Ser Ser Ala Gly Phe Leu Cys Leu His Val
35 40 45

Leu Cys Pro Ala Leu Gln Leu Thr Arg Glu Lys Tyr Lys Thr Trp Pro
50 55 60

Trp Pro Ser Phe Ile

65

<210> 459

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (21)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 459

Met Lys Glu Gly Gln Gly His Val Leu Tyr Phe Ser Arg Val Asn Cys
1 5 10 15

Lys Ala Gly His Xaa Thr Cys Arg Gln Arg Lys Pro Ala Asp Glu Leu
20 25 30

Val Cys Phe Ala Phe Gln Glu Gln Ala Pro Cys Ile Leu Leu Asn Ile
35 40 45

Arg Leu Gln Val Leu Asn Lys Tyr Leu Pro Asn Thr His Phe Leu Phe
50 55 60

Cys Val Thr Val Pro
65

<210> 460

<211> 69

<212> PRT

<213> Homo sapiens

<400> 460

Thr Met Thr Gly Ile Asp Ser Ser Pro Glu Glu Ile Leu Arg Gln Val
1 5 10 15

Gly Cys Lys Gln Gln Gln Gly Lys Gly Val Glu His Val Glu Gly Ser
20 25 30

Ser Ala Glu Ala Gly Glu Ala Ala Arg Gly Gly Gly Ala Lys Gly Gly
35 40 45

Gly Gly Ala Ala Gly Lys Gly Thr Ser Lys Val Gly Thr Leu Arg Arg
50 55 60

Thr Arg Gly Ser Thr
65

<210> 461

<211> 185

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (22)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 461

Aia	Gln	Arg	Glu	Ala	Gly	Ser	Arg	Pro	Arg	Arg	Arg	Lys	Ser	Leu	Lys
1				5					10					15	

Aia	Val	Ala	Met	Leu	Xaa	Val	Glu	Met	Gly	Gly	Gly	Cys	Arg	Gly	Ser
			20					25					30		

Met	Gly	Pro	Gly	Pro	Gly	Tyr	Ser	Ala	Gly	Ser	Arg	Val	Cys	Arg	Gly
		35					40					45			

Ser	Ser	Leu	Pro	Gln	Val	Ala	Pro	Phe	Asn	Pro	Ser	Arg	Ala	His	Leu
		50				55					60				

Leu	Pro	Pro	Pro	Val	Gly	Gly	Gly	Leu	Asn	Ser	Val	Trp	Leu	Ser	Gly
65					70					75					80

Val	Gln	Leu	Ser	Thr	Pro	Pro	Tyr	Ala	Asp	Trp	Glu	Gly	Val	Gly	Gln
				85					90					95	

Ser	Pro	Gln	Pro	Arg	Gly	Pro	Trp	Met	Gly	Ser	Ser	Ser	Leu	Gly	Thr
			100					105					110		

Val	Gly	Pro	Gly	Cys	Val	Leu	Ser	Gly	Cys	Pro	Thr	Val	Lys	Ala	Asn
		115					120					125			

Gly	Gly	Ser	Pro	Cys	Ser	Glu	Met	Leu	Gly	Glu	Arg	Arg	Leu	Leu	Glu
	130					135					140				

Pro	Ser	Val	Gly	Pro	Val	Ser	Gly	Cys	Pro	Glu	Arg	Arg	Glu	Gly	Gly
145					150					155				160	

His	Gly	Ala	Arg	Gly	Ala	Ala	Gly	Val	Val	Val	Lys	Gly	His	Ala	Ser
				165				170						175	

Val	Gln	Leu	Asn	Phe	Leu	Ser	Leu	Ile
		180					185	

<210> 462

<211> 102

<212> PRT

<213> Homo sapiens

<400> 462

Lys	Ala	Glu	Phe	Thr	Phe	Ala	Lys	Glu	Lys	Asn	Ala	Lys	Ala	Gln	Leu
1				5					10					15	

Gly	Lys	Lys	Gly	Thr	Arg	Trp	Val	Lys	His	Asp	Lys	Arg	Lys	Glu	Ile
			20					25					30		

Gln	Leu	Tyr	Gly	Cys	Val	Thr	Leu	Asn	Asp	Asp	Pro	Ser	Cys	Pro	Pro
		35					40					45			

Cys	Pro	Val	Pro	Thr	Leu	Pro	Pro	Phe	Trp	Thr	Ala	Thr	Tyr	Gly	Ser
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

50

55

60

His Gly Arg Phe Gln Lys Pro Pro Phe Ser Gln His Leu Arg Ala Gly
 65 70 75 80

Gly Ala Pro Val Gly Leu Asp Cys Gly Ala Pro Thr Gln Tyr Ala Ala
 85 90 95

Arg Pro His Gly Pro Lys
 100

<210> 463

<211> 26

<212> PRT

<213> Homo sapiens

<400> 463

Gly Cys Arg Gly Ser Met Gly Pro Gly Pro Gly Tyr Ser Ala Gly Ser
 1 5 10 15

Arg Val Cys Arg Gly Ser Ser Leu Pro Gln
 20 25

<210> 464

<211> 22

<212> PRT

<213> Homo sapiens

<400> 464

Gln Pro Arg Gly Pro Trp Met Gly Ser Ser Ser Leu Gly Thr Val Gly
 1 5 10 15

Pro Gly Cys Val Leu Ser
 20

<210> 465

<211> 21

<212> PRT

<213> Homo sapiens

<400> 465

Gly Ala Ala Gly Val Val Val Lys Gly His Ala Ser Val Gln Leu Asn
 1 5 10 15

Phe Leu Ser Leu Ile
 20

<210> 466

<211> 94

<212> PRT

<213> Homo sapiens

<400> 466

Gly Lys Pro Leu Ser Ala Ile Phe Pro Ile Cys His Met Met Phe Leu

1

10

15

Pro Gly Lys Phe Asn Leu Gly Ile Ser His Arg Cys Cys Arg Met Thr
20 25 30

Ser Pro Trp Asp Lys Arg Gln Gln Leu Arg Gln Glu Cys Lys Ser Asp
35 40 45

Pro His Val Gln Asn Pro Arg Ile His Phe Pro Glu Ser Lys Asn Ser
50 55 60

Phe Pro Ser Ala Tyr Ile Phe Val Ser Glu Gly Asn Gly Val Ser Pro
65 70 75 80

Ser Lys Trp His Cys Ile Tyr Ser Gly Thr Ser Leu Ser His
85 90

<210> 467

<211> 62

<212> PRT

<213> Homo sapiens

<400> 467

Gly Glu Arg Gly Arg Tyr Gln Ser Lys Tyr Ser Ala Thr Trp Met Val
1 5 10 15

Thr Pro His Tyr Leu Gln Thr Gln Arg Cys Lys Leu Arg Glu Met Asn
20 25 30

Ser Trp Ile Gln Gly Asn Glu Phe Leu Asp Ser Glu His Glu Gly Gln
35 40 45

Ile Tyr Ile Pro Val Ser Ile Val Asp Ala Tyr Pro Lys Asp
50 55 60

<210> 468

<211> 107

<212> PRT

<213> Homo sapiens

<400> 468

Ile Ser Ile Arg Gly Arg Ile Leu Tyr Lys Met Ala Tyr Phe Lys Val
1 5 10 15

Cys Val Ile Ile Trp Phe Gln Gln Phe Cys Val Glu Glu Thr Ser Ile
20 25 30

Ile Lys Asn Val Arg Met Leu Thr Ser Glu Phe Gln Asn Ser Tyr Ala
35 40 45

Thr Pro Val Ser Gly Leu Leu Pro Gly Ala Val Ala Trp Arg Gly Gly
50 55 60

Ala Val Tyr Gly Trp Val Arg His Ala Met Gln Val Leu Gln Lys Glu
65 70 75 80

Pro Thr Gln Pro Ser Ser Phe Leu Pro Pro Ser Asp Ala Ala Ser Phe
85 90 95

Trp Gly Pro Glu Ser Arg Leu His Leu Thr Trp
100 105

<210> 469
<211> 86
<212> PRT
<213> Homo sapiens

<400> 469
Lys Pro Phe Ala Phe Ser Ala Arg Asn Phe Pro Thr Met Leu Ser Glu
1 5 10 15

Ala Tyr Phe Gln Asp Pro Arg Met Arg Gln His His Leu Gly Val Glu
20 25 30

Arg Met Thr Val Ala Trp Val Pro Ser Ala Ile Pro Ala Trp Arg Ala
35 40 45

Ser Pro Thr Arg Thr Gln His His Pro Ser Lys Pro Gln His Gln Glu
50 55 60

Gly Ala Gln Lys Gln Gly Trp His Met Asn Ser Gly Ile Leu Met Ser
65 70 75 80

Ala Tyr Glu His Phe Leu
85

<210> 470
<211> 60
<212> PRT
<213> Homo sapiens

<400> 470
His Ser Lys Gln Asn Ile Cys Arg Glu Val Asn Ile Leu Lys Met Phe
1 5 10 15

Leu His Glu Ile Lys Lys Thr Val Thr Asp Asn Ile Ser Thr Gln Arg
20 25 30

Arg Phe Thr Tyr Asn His Gln Pro Gly Ser Val Ser Ile Phe Ser Val
35 40 45

Thr Asp Ile Leu Asp Phe Glu Val Pro Phe Gly Leu
50 55 60

<210> 471
<211> 57
<212> PRT
<213> Homo sapiens

<220>
<221> SITE

<222> (28)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 471

Lys Val Ile Asp Val Ile Phe Ser Leu Pro Pro Gly Arg Lys Ala Thr
1 5 10 15

Phe Ser Cys Pro Leu Ala Pro Leu Ser Gly Ala Xaa Gly Leu Pro Gly
20 25 30

Gly Gly Ala Asn Arg Pro Gly Pro Phe Leu Pro Cys Ile Gln Pro Trp
35 40 45

Gly Pro Leu Arg Leu Pro Glu Gly Cys
50 55

<210> 472

<211> 80

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (25)

<223> Xaa equals any of the naturally occurring L-amino acids

<400> 472

Met Ser Ser Ser Leu Cys Pro Gln Gly Gly Lys Pro Pro Ser Leu Ala
1 5 10 15

Pro Trp Pro Leu Cys Gln Gly Pro Xaa Val Cys Arg Val Gly Val Pro
20 25 30

Thr Gly Leu Ala Leu Ser Ser Pro Ala Ser Ser His Gly Gly Leu Cys
35 40 45

Asp Cys Arg Lys Val Ala Trp Leu Val Pro Gly Pro Ala Gln Ala Arg
50 55 60

Gly Arg Ala Ala Trp Phe Tyr Phe Tyr Leu Thr Leu Phe Ser Val Leu
65 70 75 80

<210> 473

<211> 26

<212> PRT

<213> Homo sapiens

<400> 473

Leu Ala Leu Ser Ser Pro Ala Ser Ser His Gly Gly Leu Cys Asp Cys
1 5 10 15

Arg Lys Val Ala Trp Leu Val Pro Gly Pro
20 25

<210> 474
 <211> 160
 <212> PRT
 <213> Homo sapiens

<220>
 <221> SITE
 <222> (124)
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 474
 Met Gln Arg Glu Arg Trp Ala Arg Pro Trp Met Ala Ser Thr Val Glu
 1 5 10 15
 Ser Arg Met Pro Glu Gly Lys Trp Arg Arg Phe Ser Thr Asp Leu Ala
 20 25 30
 Thr Trp Gly Ala Thr Pro Ala Arg Ser Trp Thr Lys Ala Ser Arg Gly
 35 40 45
 Ser Thr Thr Ala Trp Thr Arg Leu Pro Met Arg Ser Thr Met Val Leu
 50 55 60
 Asp Lys Gln Glu Arg Lys Gln Arg Ser Leu Ala Met Gly Ser Thr Thr
 65 70 75 80
 Leu Leu Asp Arg Pro Gly Arg Lys Gln Thr Lys Arg Ser Lys Gly Ser
 85 90 95
 Thr Leu Gly Ser Thr Arg Leu Gly Arg Lys Gln Arg Asn Leu Ala Lys
 100 105 110
 Gly Ser Thr Met Leu Leu Thr Arg Leu Glu Arg Xaa Trp Arg Ser Leu
 115 120 125
 Ala Gln Val Pro Thr Met Leu Leu Ala Arg Pro Gly Arg Ser Cys Arg
 130 135 140
 Met Leu Ile Met Gly Ser Thr Lys Pro Ala Arg Arg Pro Thr Ser Cys
 145 150 155 160

<210> 475
 <211> 264
 <212> PRT
 <213> Homo sapiens

<400> 475
 Met Arg Pro Leu Leu Gly Leu Leu Leu Val Phe Ala Gly Cys Thr Phe
 1 5 10 15
 Ala Leu Tyr Leu Leu Ser Thr Arg Leu Pro Arg Gly Arg Arg Leu Gly
 20 25 30

Ser Thr Glu Glu Ala Gly Gly Arg Ser Leu Trp Phe Pro Ser Asp Leu
35 40 45

Ala Glu Leu Arg Glu Leu Ser Glu Val Leu Arg Glu Tyr Arg Lys Glu
50 55 60

His Gln Ala Tyr Val Phe Leu Leu Phe Cys Gly Ala Tyr Leu Tyr Lys
65 70 75 80

Gln Gly Phe Ala Ile Pro Gly Ser Ser Phe Leu Asn Val Leu Ala Gly
85 90 95

Ala Leu Phe Gly Pro Trp Leu Gly Leu Leu Leu Cys Cys Val Leu Thr
100 105 110

Ser Val Gly Ala Thr Cys Cys Tyr Leu Leu Ser Ser Ile Phe Gly Lys
115 120 125

Gln Leu Val Val Ser Tyr Phe Pro Asp Lys Val Ala Leu Leu Gln Arg
130 135 140

Lys Val Glu Glu Asn Arg Asn Ser Leu Phe Phe Phe Leu Leu Phe Leu
145 150 155 160

Arg Leu Phe Pro Met Thr Pro Asn Trp Phe Leu Asn Leu Ser Ala Pro
165 170 175

Ile Leu Asn Ile Pro Ile Val Gln Phe Phe Phe Ser Val Leu Ile Gly
180 185 190

Leu Ile Pro Tyr Asn Phe Ile Cys Val Gln Thr Gly Ser Ile Leu Ser
195 200 205

Thr Leu Thr Ser Leu Asp Ala Leu Phe Ser Trp Asp Thr Val Phe Lys
210 215 220

Leu Leu Ala Ile Ala Met Val Ala Leu Ile Pro Gly Thr Leu Ile Lys
225 230 235 240

Lys Phe Ser Gln Lys His Leu Gln Leu Asn Glu Thr Ser Thr Ala Asn
245 250 255

His Ile His Ser Arg Lys Asp Thr
260

<210> 476

<211> 21

<212> PRT

<213> Homo sapiens

<400> 476

Asp Ile Met Pro Ala Ser Val Ile Phe Leu Ile Cys Glu Gly Val Leu
1 5 10 15

Tyr Gly Val Gln Gly
20

<210> 477
 <211> 180
 <212> PRT
 <213> Homo sapiens

<400> 477
 Gly Thr Ala Phe Gln His Ala Phe Ser Thr Asn Asp Cys Ser Arg Asn
 1 5 10 15
 Val Tyr Ile Lys Lys Asn Gly Phe Thr Leu His Arg Asn Pro Ile Ala
 20 25 30
 Gln Ser Thr Asp Gly Ala Arg Thr Lys Ile Gly Phe Ser Glu Gly Arg
 35 40 45
 His Ala Trp Glu Val Trp Trp Glu Gly Pro Leu Gly Thr Val Ala Val
 50 55 60
 Ile Gly Ile Ala Thr Lys Arg Ala Pro Met Gln Cys Gln Gly Tyr Val
 65 70 75 80
 Ala Leu Leu Gly Ser Asp Asp Gln Ser Trp Gly Trp Asn Leu Val Asp
 85 90 95
 Asn Asn Leu Leu His Asn Gly Glu Val Asn Gly Ser Phe Pro Gln Cys
 100 105 110
 Asn Asn Ala Pro Lys Tyr Gln Ile Gly Glu Arg Ile Arg Val Ile Leu
 115 120 125
 Asp Met Glu Asp Lys Thr Leu Ala Phe Glu Arg Gly Tyr Glu Phe Leu
 130 135 140
 Gly Val Ala Phe Arg Gly Leu Pro Lys Val Cys Leu Tyr Pro Ala Val
 145 150 155 160
 Ser Ala Val Tyr Gly Asn Thr Glu Val Thr Leu Val Tyr Leu Gly Lys
 165 170 175
 Pro Leu Asp Gly
 180

<210> 478
 <211> 35
 <212> PRT
 <213> Homo sapiens

<400> 478
 Ala Arg Ala Phe Gln His Leu Met Val Ala Asp His Ser His Phe His
 1 5 10 15
 Arg Thr Leu Ile Lys Gln Pro Ser Met Ile Pro Asn Ala Thr Phe Tyr
 20 25 30
 His Ile Phe

35

<210> 479
<211> 131
<212> PRT
<213> Homo sapiens

<400> 479
Ala Arg Ala Leu Pro Glu Ile Lys Gly Ser Arg Leu Gln Glu Ile Asn
1 5 10 15
Asp Val Cys Ala Ile Cys Tyr His Glu Phe Thr Thr Ser Ala Arg Ile
20 25 30
Thr Pro Cys Asn His Tyr Phe His Ala Leu Cys Leu Arg Lys Trp Leu
35 40 45
Tyr Ile Gln Asp Thr Cys Pro Met Cys His Gln Lys Val Tyr Ile Glu
50 55 60
Asp Asp Ile Lys Asp Asn Ser Asn Val Ser Asn Asn Asn Gly Phe Ile
65 70 75 80
Pro Pro Asn Glu Thr Pro Glu Glu Ala Val Arg Glu Ala Ala Glu
85 90 95
Ser Asp Arg Glu Leu Asn Glu Asp Asp Ser Thr Asp Cys Asp Asp Asp
100 105 110
Val Gln Arg Glu Arg Asn Gly Val Ile Gln His Thr Gly Ala Ala Ala
115 120 125
Gly Arg Ile
130

<210> 480
<211> 16
<212> PRT
<213> Homo sapiens

<400> 480
Phe Ser Thr Gln Ala Gln Gln Leu Glu Glu Phe Asn Asp Asp Thr Asp
1 5 10 15

<210> 481
<211> 22
<212> PRT
<213> Homo sapiens

<400> 481
Arg Leu Gln Glu Ile Asn Asp Val Cys Ala Ile Cys Tyr His Glu Phe
1 5 10 15

Thr Thr Ser Ala Arg Ile
20

<210> 482
<211> 20
<212> PRT
<213> Homo sapiens

<400> 482
Leu Tyr Ile Gln Asp Thr Cys Pro Met Cys His Gln Lys Val Tyr Ile
1 5 10 15

Glu Asp Asp Ile
20

<210> 483
<211> 21
<212> PRT
<213> Homo sapiens

<400> 483
Val Ser Asn Asn Asn Gly Phe Ile Pro Pro Asn Glu Thr Pro Glu Glu
1 5 10 15

Ala Val Arg Glu Ala
20

<210> 484
<211> 26
<212> PRT
<213> Homo sapiens

<400> 484
Asp Asp Ser Thr Asp Cys Asp Asp Asp Val Gln Arg Glu Arg Asn Gly
1 5 10 15

Val Ile Gln His Thr Gly Ala Ala Ala Gly
20 25

<210> 485
<211> 141
<212> PRT
<213> Homo sapiens

<220>
<221> SITE
<222> (54)
<223> Xaa equals any of the naturally occurring L-amino acids

<400> 485
Val Ala Gly Ile Thr Gly Ala His His His Ala Gln Leu Ile Phe Val
1 5 10 15

Leu Leu Val Glu Met Gly Phe His His Val Gly Gln Ala Gly Leu Lys
20 25 30

Leu Leu Thr Ser Asp Asn Pro Arg Thr Ser Ala Ser Gln Ser Ala Gly
35 40 45

Ile Thr Gly Met Ser Xaa Gly Arg Arg Ile Thr Cys Gly Gln Glu Phe
50 55 60

Lys Thr Ala Val Ser Tyr Asn Cys Thr Thr Ala Leu Gln Pro Asp Arg
65 70 75 80

Ala Lys Leu Cys Phe Leu Phe Lys Lys Lys Lys Lys Ile Ser Ile Gln
85 90 95

Arg Thr Leu Pro Gly Ile Lys Arg Val Ile Tyr Asn Tyr Glu Arg Val
100 105 110

Asp Ser Ser Lys Gly His Asn Ser Gln Val Gln Trp Ala His Ala Cys
115 120 125

Asn Pro Ser Thr Leu Gly Gly Arg Gly Gly Gln Ile Val
130 135 140

<210> 486

<211> 22

<212> PRT

<213> Homo sapiens

<400> 486

Ala Gly Ile Thr Gly Ala His His His Ala Gln Leu Ile Phe Val Leu
1 5 10 15

Leu Val Glu Met Gly Phe
20

<210> 487

<211> 27

<212> PRT

<213> Homo sapiens

<400> 487

Arg Val Ile Tyr Asn Tyr Glu Arg Val Asp Ser Ser Lys Gly His Asn
1 5 10 15

Ser Gln Val Gln Trp Ala His Ala Cys Asn Pro
20 25

<210> 488

<211> 106

<212> PRT

<213> Homo sapiens

<400> 488

Ala Gly Ala Glu Val Val Met Leu Phe Leu Leu Thr Pro Ser Ser His

1

5

10

15

His Gln His Glu Cys Val Arg Arg Ala Phe Glu Cys Gly Asp Cys His
20 25 30

Ile Leu Leu Asp Asn Asn Val Leu Gly Val Asp Cys His Gly Ala Gly
35 40 45

Glu Arg Ala Val His Leu Glu Asp His Phe Val His Ile Asp Thr Ile
50 55 60

Ser Leu Leu Leu Glu Asp Ala Leu Glu Tyr Ser Ala Leu Ile Ala Gly
65 70 75 80

His Pro Lys Ser Asp Leu Pro Pro Gly Leu Ser Arg Cys Arg Pro Trp
85 90 95

Glu His His Trp Pro Ile Ser Tyr Thr Gly
100 105

<210> 489

<211> 64

<212> PRT

<213> Homo sapiens

<400> 489

Thr Ile Ser Tyr Leu Cys Asn Asn Val Ser Tyr Met Gln Leu Gln Lys
1 5 10 15

Leu Val Gly Lys Ser Met Ile Phe Leu Pro Tyr Ser Leu Pro Ile His
20 25 30

Leu Pro Gly Asn His Arg Leu Leu Leu Pro Arg Val Gly Met Arg Leu
35 40 45

Arg Gly Cys Cys Phe Ser Pro Tyr Ile Ile Thr Asp Phe Lys Trp Cys
50 55 60

<210> 490

<211> 58

<212> PRT

<213> Homo sapiens

<400> 490

Glu Met Gly Gln Trp Cys Ser Gln Gly Leu His Leu Asp Ser Pro Gly
1 5 10 15

Gly Lys Ser Asp Phe Gly Cys Pro Ala Ile Asn Ala Glu Tyr Ser Arg
20 25 30

Ala Ser Ser Lys Ser Arg Leu Met Val Ser Met Trp Thr Lys Trp Ser
35 40 45

Ser Arg Cys Thr Ala Leu Ser Pro Ala Pro
50 55

<210> 491
<211> 25
<212> PRT
<213> Homo sapiens

<400> 491
Arg Ala Phe Glu Cys Gly Asp Cys His Ile Leu Leu Asp Asn Asn Val
1 5 10 15

Leu Gly Val Asp Cys His Gly Ala Gly
20 25

<210> 492
<211> 23
<212> PRT
<213> Homo sapiens

<400> 492
Leu Val Gly Lys Ser Met Ile Phe Leu Pro Tyr Ser Leu Pro Ile His
1 5 10 15

Leu Pro Gly Asn His Arg Leu
20



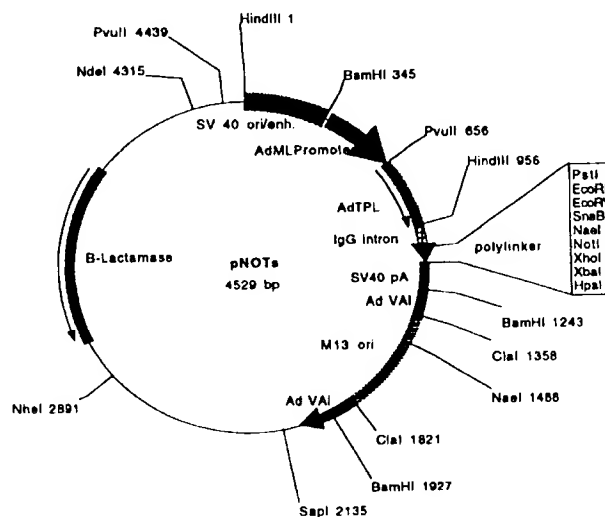
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 5/10, C07K 14/47, C12Q 1/68, A61K 38/17		A2	(11) International Publication Number: WO 98/30696
(21) International Application Number: PCT/US98/00575		(43) International Publication Date: 16 July 1998 (16.07.98)	
(22) International Filing Date: 13 January 1998 (13.01.98)		(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).	
(30) Priority Data: 08/783,395 13 January 1997 (13.01.97) AP (34) Countries for which the regional or international application was filed: PCT/US97/05682 4 April 1997 (04.04.97) WO (34) Countries for which the regional or international application was filed: JP et al. 08/866,022 30 May 1997 (30.05.97) US 08/924,838 5 September 1997 (05.09.97) US 09/005,986 12 January 1998 (12.01.98) US		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).		Published Without international search report and to be republished upon receipt of that report.	
(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).			

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al. 1989. Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

5

10

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

15

This application claims the priority of the following applications: (1) Ser. No. 08/866,022, filed May 30, 1997, which is a continuation-in-part of Ser. No. 08/628,364, filed April 5, 1996, now abandoned; (2) Ser. No. 08/XXX,XXX, filed January 12, 1998, which is a continuation-in-part of Ser. No. 08/924,838, filed September 5, 1997, which is
20 a divisional of Ser. No. 08/628,364, filed April 5, 1996, now abandoned; (3) Ser. No. 08/924,838, filed September 5, 1997, which is a divisional of Ser. No. 08/628,364, filed April 5, 1996, now abandoned; (4) Ser. No. 08/783,395, filed January 13, 1997, which is a continuation-in-part of Ser. No. 08/628,364, filed April 5, 1996, now abandoned; and (5)
25 international application Ser. No. PCT/US97/05682, filed April 4, 1997, claiming the priority of Ser. No. 08/628,364, filed April 5, 1996, now abandoned; all of which are incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by
30 such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines,
35 such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression
40 cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader

sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 370 to nucleotide 1341;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 700 to nucleotide 1341;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 370 to nucleotide 798;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone B121 deposited under accession number ATCC 98019;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone B121 deposited under accession number ATCC 98019;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone B121 deposited under accession number ATCC 98019;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone B121 deposited under accession number ATCC 98019;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

5 Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 370 to nucleotide 1341; the nucleotide sequence of SEQ ID NO:1 from nucleotide 700 to nucleotide 1341; the nucleotide sequence of SEQ ID NO:1 from nucleotide 370 to nucleotide 798; the nucleotide sequence of the full-length protein coding sequence of clone B121 deposited under accession number ATCC 98019; or the nucleotide
10 sequence of the mature protein coding sequence of clone B121 deposited under accession number ATCC 98019. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone B121 deposited under accession number ATCC 98019. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid
15 sequence of SEQ ID NO:2 from amino acid 1 to amino acid 143.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group
20 consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 143;

(c) fragments of the amino acid sequence of SEQ ID NO:2; and

25 (d) the amino acid sequence encoded by the cDNA insert of clone B121 deposited under accession number ATCC 98019;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 143.

30 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 24 to nucleotide 548;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 525 to nucleotide 548;

5 (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 115 to nucleotide 317;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone B196 deposited under accession number ATCC 98021;

10 (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone B196 deposited under accession number ATCC 98021;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone B196 deposited under accession number ATCC 98021;

15 (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone B196 deposited under accession number ATCC 98021;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

20 (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

25 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 24 to nucleotide 548; the nucleotide sequence of SEQ ID NO:3 from nucleotide 525 to nucleotide 548; the nucleotide sequence of SEQ ID NO:3 from
30 nucleotide 115 to nucleotide 317; the nucleotide sequence of the full-length protein coding sequence of clone B196 deposited under accession number ATCC 98021; or the nucleotide sequence of the mature protein coding sequence of clone B196 deposited under accession number ATCC 98021. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone B196 deposited under

accession number ATCC 98021. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 118 to amino acid 162.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3 or SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
 - (b) the amino acid sequence of SEQ ID NO:4 from amino acid 118 to amino acid 162;
 - (c) fragments of the amino acid sequence of SEQ ID NO:4; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone B196 deposited under accession number ATCC 98021;
- the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 118 to amino acid 162.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 50 to nucleotide 538;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 290 to nucleotide 538;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone D157 deposited under accession number ATCC 98020;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone D157 deposited under accession number ATCC 98020;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone D157 deposited under accession number ATCC 98020;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone D157 deposited under accession number ATCC 98020;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7;

5 (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:7 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

10 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:6 from nucleotide 50 to nucleotide 538; the nucleotide sequence of SEQ ID NO:6 from nucleotide 290 to nucleotide 538; the nucleotide sequence of the full-length protein coding sequence of clone D157 deposited under accession number ATCC 98020; or the nucleotide sequence of the mature protein coding sequence of clone D157 deposited under accession number ATCC 98020. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone D157 deposited under accession number ATCC 98020.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:6.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:7;

(b) fragments of the amino acid sequence of SEQ ID NO:7; and

(c) the amino acid sequence encoded by the cDNA insert of clone D157 deposited under accession number ATCC 98020;

30 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:7.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions.

Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

- 5 (a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and
- (b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

20 Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

Fig. 2 is an autoradiograph evidencing the expression of clone B121 in baculovirus (bands of expressed protein are indicated by dots).

Fig. 3 is an autoradiograph evidencing the expression of clone B121 in baculovirus (bands of expressed protein are indicated by dots).

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by

expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

5 As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins
10 which are transported across the membrane of the endoplasmic reticulum.

Clone "B121"

A polynucleotide of the present invention has been identified as clone "B121". B121 was isolated from a human adult blood (peripheral blood mononuclear cells treated
15 with concanavalin-A and phorbol myristate acetate) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. B121 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "B121
20 protein").

The nucleotide sequence of B121 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the B121 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 98 to 110 are a predicted leader/signal
25 sequence, with the predicted mature amino acid sequence beginning at amino acid 111, or are a transmembrane domain.

Clone B121 was deposited on April 4, 1996 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98019. All restrictions on the availability to the public of the deposited
30 material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b). The clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site NotI). The EcoRI/NotI restriction fragment obtainable from the deposit containing clone B121 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for B121 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. B121 demonstrated at least some identity with sequences identified as R83586 (yp16a07.r1 Homo sapiens cDNA clone 187572 5'), H23221 (ym52f07.s1 Homo sapiens cDNA clone 51884 3'), W72694 (zd68f10.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 345835 3' similar to contains Alu repetitive element), and AA136867 (zl01c02.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 491042 3'). The predicted amino acid sequence disclosed herein for B121 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted B121 protein demonstrated at least some identity with sequences identified as U28928 (C44B7.4 gene product [Caenorhabditis elegans]). Based upon identity, B121 proteins and each identical protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the B121 protein sequence, one around amino acid 70 and another around amino acid 200 of SEQ ID NO:2.

Figures 2 and 3 are autoradiographs evidencing expression of clone B121 of the present invention. Clone B121 was expressed in baculovirus; dots indicate the bands of expressed B121 protein not present in the control lanes ("mock").

Clone "B196"

A polynucleotide of the present invention has been identified as clone "B196". B196 was isolated from a human adult blood (peripheral blood mononuclear cell) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. B196 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "B196 protein").

The nucleotide sequence of B196 as presently determined is reported in SEQ ID NO:1. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:2. The predicted amino acid sequence of the B196 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 155 to 167 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 168, or are a transmembrane domain. Additional nucleotide sequence for B196 is reported in SEQ ID NO:3. Applicants believe SEQ ID

NO:3 represents a cDNA molecule produced from an immature mRNA transcript, as base pairs 205 to 352 of SEQ ID NO:3 appear to be an intron sequence. SEQ ID NO:1 was derived from SEQ ID NO:3 by deleting this presumed intron sequence.

Clone B196 was deposited on April 4, 1996 with the American Type Culture
5 Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98021. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b). The clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site
10 NotI). The EcoRI/NotI restriction fragment obtainable from the deposit containing clone B196 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for B196 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. B196 demonstrated at least some similarity with sequences
15 identified as AA235452 (zt35c01.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 724320 3' similar to contains Alu repetitive element), T09157 (EST07050 Homo sapiens cDNA clone HIBBP87 5' end), T34456 (EST68380 Homo sapiens cDNA 5' end similar to None), T35039 (EST79238 Homo sapiens cDNA similar to None), and T70971 (yc49f08.r1 Homo sapiens cDNA clone 84039 5'). Based upon sequence similarity, B196
20 proteins and each similar protein or peptide may share at least some activity.

Clone "D157"

A polynucleotide of the present invention has been identified as clone "D157". D157 was isolated from a human adult blood (peripheral blood mononuclear cell) cDNA
25 library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. D157 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "D157 protein").

30 The nucleotide sequence of D157 as presently determined is reported in SEQ ID NO:6. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the D157 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:7. Amino acids 1 to 80 are a predicted leader/signal

sequence, with the predicted mature amino acid sequence beginning at amino acid 81, or are a transmembrane domain.

Clone D157 was deposited on April 4, 1996 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98020. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b). The clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site NotI).

The nucleotide sequence disclosed herein for D157 was searched against the GenBank database using BLASTA/BLASTX and FASTA search protocols. D157 demonstrated at least some identity with an EST identified as "yd93h05.s1 Homo sapiens cDNA clone 115833 3'" found at GenBank accession number T87909, an EST identified as "yo72g04.r1 Homo sapiens cDNA clone 183510 5'" found at GenBank accession number H45571, and an EST identified as "yo72g03.s1 Homo sapiens cDNA clone 183508 3'" found at GenBank accession number H45474. D157 and these ESTs demonstrate some homology or similarity with the sequence for human peripheral myelin protein 22 (GenBank accession number D11428). Based upon identity, D157 proteins and each identical protein or peptide may share at least some activity.

20

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, *Bio/Technology* 10, 773-778 (1992) and in R.S. McDowell, *et al.*, *J. Amer. Chem. Soc.* 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

30

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence

listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15**(7): 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62**(1): 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through

deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14**(9): 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90**(16): 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91**(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614,396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or

polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

5 The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides .

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

10 The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as
15 stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) [†]	Hybridization Temperature and Buffer [‡]	Wash Temperature and Buffer [‡]
5	A	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
	B	<50	T _B *; 1xSSC	T _B *; 1xSSC
	C	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
	D	<50	T _D *; 1xSSC	T _D *; 1xSSC
	E	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
	F	<50	T _F *; 1xSSC	T _F *; 1xSSC
10	G	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
	H	<50	T _H *; 4xSSC	T _H *; 4xSSC
	I	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
	J	<50	T _J *; 4xSSC	T _J *; 4xSSC
	K	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
	L	<50	T _L *; 2xSSC	T _L *; 2xSSC
15	M	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
	N	<50	T _N *; 6xSSC	T _N *; 6xSSC
	O	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
	P	<50	T _P *; 6xSSC	T _P *; 6xSSC
	Q	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
	R	<50	T _R *; 4xSSC	T _R *; 4xSSC

[†]: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

[‡]: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH₂PO₄, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6(log₁₀[Na⁺]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na⁺] is the concentration of sodium ions in the hybridization buffer ([Na⁺] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds.,

5 John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or
10 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an
15 expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably
20 linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the
25 protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

30 Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant

methy1 or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance
5 with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

10 The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith,
15 including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally
20 provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another
25 amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be
30 expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays

for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured
5 by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter
10 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology* 133:327-341, 1991; Bertagnolli, et al., *J. Immunol.* 149:3778-3783, 1992; Bowman et al., *J. Immunol.* 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node
15 cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ , Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

20 Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., *J. Exp. Med.* 173:1205-1211, 1991; Moreau et al., *Nature*
25 336:690-692, 1988; Greenberger et al., *Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991;
30 Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient
5 by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic
10 acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function
15 (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides.
20 For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used
25 to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II
30 molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowman et al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*, J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. *Cellular Biology* 15:141-151, 1995; Keller et al., *Molecular and Cellular Biology* 13:473-486, 1993; McClanahan et al., *Blood* 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., *Proc. Natl. Acad. Sci. USA* 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., *Experimental Hematology* 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

5 A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

10 A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De*
15 *nov*o bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce
20 differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

25 Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and
30 other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. *De novo* tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce
5 differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in
10 the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve
15 tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present
20 invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of
25 non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac)
30 and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

5 A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

10 Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

15 Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

20 A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin- β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, 30 United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

5

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells.

10 Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses
15 against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population
20 of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent
25 chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene
30 Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

10 The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 15 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

30 The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 5 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in 10 the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat 15 inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting 20 from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major 25 roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

30 The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

5 E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to
10 their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention
15 encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue
20 in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and
25 polynucleotides of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the
30 cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

10 Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

20 Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein

and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines

or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If
5 administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical
10 composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is
15 administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention.
20 When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid
25 form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present
30 invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium

Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

5 The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician
10 will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μ g to about 100
15 mg (preferably about 0.1mg to about 10 mg, more preferably about 0.1 μ g to about 1 mg) of protein of the present invention per kg body weight.

 The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is
20 contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

 Protein of the invention may also be used to immunize animals to obtain
25 polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in
30 R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where

abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

5 For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably
10 be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the
15 methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical
20 applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium
25 sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other
30 ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbition of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- α and TGF- β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect

the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

5 Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

10 Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Agostino, Michael J.
- (ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
- (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Genetics Institute, Inc.
 - (B) STREET: 87 CambridgePark Drive
 - (C) CITY: Cambridge
 - (D) STATE: MA
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 02140
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Sprunger, Suzanne A.
 - (B) REGISTRATION NUMBER: 41,323
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (617) 498-8284
 - (B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1731 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCCGGCAGGT TGCCAGCGTC GCTACAGCCC AGACCAAGGC AGAATAATCT CCGGATGAGC 60
TGGTGGCACC GCTGAGCCTT TGGTCTCACC AGGGCTTCCT GTTGCTGGCA GGCGGGGTGG 120
AGCGGAGCTG CTGGGAGGCT GCTGGATAGG AGAGGGGTCA CGGCTGCGGA AGAGGAGGTT 180
CTTCGGGACA CCCGTGGATG GACACGGCAA GGAAACACCA GGCCAACCAC AGCTGGGGAT 240
AAAATAGCAC AACCACACCC TGCCGTCCAG CGCCTCCCAG CCTGTGCCCC TTCCTAGTAC 300
CACCAGCAAC CATCAATCCC GTCTCCTCCT GCCTCCTCTC CTGCAATCCA CCGCGCCACG 360
ACTATCGCCA TGGCAGCCCT GATCGCAGAG AACTTCCGCT TCCTGTCACT TTTCTTCAAG 420
AGCAAGGATG TGATGATTTT CAACGGCCTG GTGGCACTGG GCACGGTGGG CAGCCAGGAG 480
CTGTCCTCTG TGGTGGCCTT CCACTGCCCC TGCTCGCCGG CCCGGAACTA CCTGTACGGG 540
CTGGCGGCCA TCGGCGTGCC CGCCCTGGTG CTCTTCATCA TTGGCATCAT CCTCAACAAC 600
CACACCTGGA ACCTCGTGGG CGAGTGCCAG CACCGGAGGA CCAAGAACTG CTCCGCCGCC 660
CCCACCTTCC TCCTTCTAAG CTCCATCCTG GGACGTGCGG CTGTGGCCCC TGTCACCTGG 720
TCTGTCATCT CCCTGCTGCG TGGTGAGGCT TATGTCTGTG CTCTCAGTGA GTTCGTGGAC 780
CCTTCTCAC TCACGGCCAG GGAAGAGCAC TTCCCATCAG CCCACGCCAC TGAATCCTG 840
GCCAGGTTC CCTGCAAGGA GAACCCTGAC AACCTGTCAG ACTTCCGGA GGAGGTCAGC 900
CGCAGGCTCA GGTATGAGTC CCAGCTCTTT GGATGGCTGC TCATCGGCGT GGTGGCCATC 960
CTGGTGTTCC TGACCAAGTG CCTCAAGCAT TACTGCTCAC CACTCAGCTA CCGCCAGGAG 1020
GCCTACTGGG CGCAGTACCG CGCCAATGAG GACCAGCTGT TCCAGCGCAC GGCCGAGGTG 1080
CACTCTCGGG TGCTCGCTGC CAACAATGTG CGCCGCTTCT TTGGCTTTGT GGCGCTCAAC 1140
AAGGATGATG AGGAACTGAT TGCCAACTTC CCAGTGGAAG GCACGCAGCC ACGGCCACAG 1200
TGGAATGCCA TCACCGGCGT CTACTTGTAAC CGTGAGAACC AGGGCCTCCC ACTCTACAGC 1260
CGCCTGCACA AGTGGGCCCA GGGTCTGGCA GGCAACGGCG CGGCCCCCTGA CAACGTGGAG 1320
ATGGCCCTGC TCCCCTCCTA AGGAGGTGCT TCCCATGCTC TTTGTAAATG GCACTGCTTG 1380
GTCCCAAAC TGAACCCACT GCTTGCTCAC ATCCATATCA GAAGGGGATT TTAAAAAAC 1440
TGTTATCTTC TTGGCCAGGG GAAAGGACCA CAAGGCAATC TGGGGTGTGG ACAGACCCAG 1500
TAGACAATGG AAGCCCCAGC CAGCAGGGCC AGGTGACAGT GAAGCTCACC AGTGGGCTCC 1560

TTTATGGTAC TCTATGCAGT TAACATGTAT CTAGCTGCAT AGGGACACCC AGCGCAGCAG 1620
 TGCACCACTG GGAAGTGGCC TCCAGTGCAG CCTCTGGCCT TATTTTATAT ATTTAAATTT 1680
 TTGATAAAGT TTTTCTTACT AAAAGGAAAA AAAAAAAAAA AAAAAAAAAA A 1731

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 323 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ala	Ala	Leu	Ile	Ala	Glu	Asn	Phe	Arg	Phe	Leu	Ser	Leu	Phe	Phe	1	5	10	15
Lys	Ser	Lys	Asp	Val	Met	Ile	Phe	Asn	Gly	Leu	Val	Ala	Leu	Gly	Thr	20	25	30	
Val	Gly	Ser	Gln	Glu	Leu	Ser	Ser	Val	Val	Ala	Phe	His	Cys	Pro	Cys	35	40	45	
Ser	Pro	Ala	Arg	Asn	Tyr	Leu	Tyr	Gly	Leu	Ala	Ala	Ile	Gly	Val	Pro	50	55	60	
Ala	Leu	Val	Leu	Phe	Ile	Ile	Gly	Ile	Ile	Leu	Asn	Asn	His	Thr	Trp	65	70	75	80
Asn	Leu	Val	Gly	Glu	Cys	Gln	His	Arg	Arg	Thr	Lys	Asn	Cys	Ser	Ala	85	90	95	
Ala	Pro	Thr	Phe	Leu	Leu	Leu	Ser	Ser	Ile	Leu	Gly	Arg	Ala	Ala	Val	100	105	110	
Ala	Pro	Val	Thr	Trp	Ser	Val	Ile	Ser	Leu	Leu	Arg	Gly	Glu	Ala	Tyr	115	120	125	
Val	Cys	Ala	Leu	Ser	Glu	Phe	Val	Asp	Pro	Ser	Ser	Leu	Thr	Ala	Arg	130	135	140	
Glu	Glu	His	Phe	Pro	Ser	Ala	His	Ala	Thr	Glu	Ile	Leu	Ala	Arg	Phe	145	150	155	160
Pro	Cys	Lys	Glu	Asn	Pro	Asp	Asn	Leu	Ser	Asp	Phe	Arg	Glu	Glu	Val	165	170	175	
Ser	Arg	Arg	Leu	Arg	Tyr	Glu	Ser	Gln	Leu	Phe	Gly	Trp	Leu	Leu	Ile				

180

185

190

Gly Val Val Ala Ile Leu Val Phe Leu Thr Lys Cys Leu Lys His Tyr
 195 200 205
 Cys Ser Pro Leu Ser Tyr Arg Gln Glu Ala Tyr Trp Ala Gln Tyr Arg
 210 215 220
 Ala Asn Glu Asp Gln Leu Phe Gln Arg Thr Ala Glu Val His Ser Arg
 225 230 235 240
 Val Leu Ala Ala Asn Asn Val Arg Arg Phe Phe Gly Phe Val Ala Leu
 245 250 255
 Asn Lys Asp Asp Glu Glu Leu Ile Ala Asn Phe Pro Val Glu Gly Thr
 260 265 270
 Gln Pro Arg Pro Gln Trp Asn Ala Ile Thr Gly Val Tyr Leu Tyr Arg
 275 280 285
 Glu Asn Gln Gly Leu Pro Leu Tyr Ser Arg Leu His Lys Trp Ala Gln
 290 295 300
 Gly Leu Ala Gly Asn Gly Ala Ala Pro Asp Asn Val Glu Met Ala Leu
 305 310 315 320
 Leu Pro Ser

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1522 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCCGACGGC GGCCGAGACG GACATGAAGC AATATCAAGG CTCCGGCGGC GTCGCCATGG 60
 ATGTGGAACG GAGTCGCTTC CCCTACTGCG TGGTGTGGAC GCCCATCCCG GTGCTCACGT 120
 GGTTTTCCC CATCATCGGC CACATGGGCA TCTGCACATC CACAGGAGTC ATTCGGGACT 180
 TCGCGGGCCC CTA CTTTGTGTC TCAGAGGACA ACATGGCCTT TGGAAAGCCT GCCAAGTACT 240
 GGAAGTTGGA CCCTGCTCAG GTCTATGCTA GCGGGCCCAA CGCATGGGAC ACGGCTGTGC 300
 ACGACGCCTC TGAGGAGTAC AAGCACCGCA TGCACAATCT CTGCTGTGAC AACTGCCACT 360

CGCACGTGGC ATYGGCCCTG AATCTGATGC GCTACAACAA CAGCACCAAC TGGAATATGG 420
 TGACGCTCTG CTTCTTCTGC CTGCTCTACG GGAAGTACGT CAGCGTTGGG GCCTTCGTGA 480
 AGACCTGGCT GCCCTTCATC CTTCTCCTGG GCATCATCCT CACCGTCAGC CTGGTCTTTA 540
 ACCTCCGGTG ATGGCTGCTC GGTGGCCCCA CACCCACCAG GTTCCCGAGG AAACAGCCGC 600
 CATCCCTTTT GGTTCAGAT TTTTTTCTCC TCACCCCAA AGGCAGGGTT GGGCCTGCTG 660
 TTGTGGACCG GGGGTCGGGG CTGGCAGGAT GGAAGGACTG AGGACCAGCA TGAAGTGGGG 720
 GTTTGTTGTC TCCCTGCCTC TCAGAAGCAC CCTGTCCCCT CCTCCCCAGG CCTGTGACTC 780
 CGGCCCTGGA AGCCCCTTTG TTCTTCTGTT GAAAGGCTTT GGCTTCCCTC TGTAAGAGCTG 840
 CTCCCGCCAC CACCTGCTGG GGTCTGCCT CAGCCCAGTG CCCAGTATGG GGAGAGGAGG 900
 ACATTTGGGC TCACCTGTCA AGGTGGCCCT GGGACCAGAG CTGGTCCCAR CATGGGGTGC 960
 ACCGGGTACA CTTAACGTGT CTCTATAARC CAAGTTGCTT CAGGACCTTC ACCACTGGCC 1020
 TCTAGAATGG TCCAGAGGGG CTGGCTGGGT CCCTTTGTMA GACTCCTGCC GGCAGCTKCC 1080
 CTGGGGGACA TGTGTGCCCC TCTGGCATCC TCCAGCCCGT GCAGTCCGCT CTTCACTGTT 1140
 CCACGGCCTC CCAGTGCCTC CCAGCATTGG ACCCATCTCC CCCTGCAGTT TGAGGCCAGA 1200
 GAGGTGAGTG GACCTGACAA GTGCCAGAGT AACCCTGTAG ACAGAGCAGT GTAGACAGCG 1260
 CTCAGCCCCA GCCCCAGGTG TGGACCTCAT GCTGGTGATG GCTCCCCTGG GTGGCCTGCC 1320
 AGCACAGCCA GTKCCATCAG GGAGCTGAAG GGGCTGTCCC CCACCTAACT CCAGCTCCCC 1380
 CTTACGTTG TCACCAAGGC CCTGTGCCGC CCGCCTCGCC CCCCTGCTCT GTGGATTCCT 1440
 TTGGGAAGGG CTCCC'TGGGC AGGACAATAA AGAGTTTTGA CTCCAAAAAA AAAAAAAAAA 1500
 AAAAAAAAAA AAAAAAAAAA AA 1522

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 175 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Lys Gln Tyr Gln Gly Ser Gly Gly Val Ala Met Asp Val Glu Arg
 1 5 10 15
 Ser Arg Phe Pro Tyr Cys Val Val Trp Thr Pro Ile Pro Val Leu Thr
 20 25 30
 Trp Phe Phe Pro Ile Ile Gly His Met Gly Ile Cys Thr Ser Thr Gly
 35 40 45
 Val Ile Arg Asp Phe Ala Gly Pro Tyr Phe Val Ser Glu Asp Asn Met
 50 55 60
 Ala Phe Gly Lys Pro Ala Lys Tyr Trp Lys Leu Asp Pro Ala Gln Val
 65 70 75 80
 Tyr Ala Ser Gly Pro Asn Ala Trp Asp Thr Ala Val His Asp Ala Ser
 85 90 95
 Glu Glu Tyr Lys His Arg Met His Asn Leu Cys Cys Asp Asn Cys His
 100 105 110
 Ser His Val Ala Xaa Ala Leu Asn Leu Met Arg Tyr Asn Asn Ser Thr
 115 120 125
 Asn Trp Asn Met Val Thr Leu Cys Phe Phe Cys Leu Leu Tyr Gly Lys
 130 135 140
 Tyr Val Ser Val Gly Ala Phe Val Lys Thr Trp Leu Pro Phe Ile Leu
 145 150 155 160
 Leu Leu Gly Ile Ile Leu Thr Val Ser Leu Val Phe Asn Leu Arg
 165 170 175

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1670 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGCCGACGGC GGCCGAGACG GACATGAAGC AATATCAAGG CTCCGGCGGC GTCGCCATGG 60
 ATGTGGAACG GAGTCGCTTC CCCTACTGCG TGGTGTGGAC GCCCATCCCG GTGCTCACGT 120
 GGTTTTCCC CATCATCGGC CACATGGGCA TCTGCACATC CACAGGAGTC ATTCGGGACT 180
 TCGCGGGCCC CTACTTTGTC TCAGCCGGGA GGGCAGTGTG GCCAGAAGGA TTCTTAAGTA 240

ACTGACCCAG CCCTTTGCCC CCACCCCTGG GGTACCGAGA CATGGGTAGG GATTAGAGGC 300
AAGAGTGGAG AGTCAGACCA TCCAGGAACC ACATCTCTGG ACCTTCAGAA GGAGGACAAC 360
ATGGCCTTTG GAAAGCCTGC CAAGTACTGG AAGTTGGACC CTGCTCAGGT CTATGCTAGC 420
GGGCCCCAAG CATGGGACAC GGCTGTGCAC GACGCCTCTG AGGAGTACAA GCACCGCATG 480
CACAATCTCT GCTGTGACAA CTGCCACTCG CACGTGGCAT YGGCCCTGAA TCTGATGCGC 540
TACAACAACA GCACCAACTG GAATATGGTG ACGCTCTGCT TCTTCTGCCT GCTCTACGGG 600
AAGTACGTCA GCGTTGGGGC CTTCTGTAAG ACCTGGCTGC CCTTCATCCT TCTCCTGGGC 660
ATCATCCTCA CCGTCAGCCT GGTCTTTAAC CTCCGGTGAT GGCTGCTCGG TGGCCCCACA 720
CCCACCAGGG TCCCGAGGAA ACAGCCGCCA TCCCTTTTGG TTCCAGATTT TTTTCTCCTC 780
ACCCCAAAG GCAGGGTTGG GCCTGCTGTT GTGGACCGGG GGTGCGGGCT GGCAGGATGG 840
AAGGACTGAG GACCAGCATG AAGTGGGGGT TTGTTGTCTC CCTGCCTCTC AGAAGCACCC 900
TGTCCCCTCC TCCCAGGCC TGTGACTCCG GCCCTGGAAG CCCCTTTGTT CTTCTGTTGA 960
AAGGCTTTGG CTTCCCTCTG TAGAGCTGCT CCCGCCACCA CCTGCTGGGG TCCTGCCTCA 1020
GCCCAGTGCC CAGTATGGGG AGAGGAGGAC ATTTGGGCTC ACCTGTCAAG GTGGCCCTGG 1080
GACCAGAGCT GGTCCCARCA TGGGGTGCAC CGGGTACACT TAACGTGTCT CTATAARCCA 1140
AGTTGCTTCA GGACCTTCAC CACTGGCCTC TAGAATGGTC CAGAGGGGCT GGCTGGGTCC 1200
CTTTGTMAGA CTCTGCCGG CAGCTKCCCT GGGGGACATG TGTGCCCATC TGGCATCCTC 1260
CAGCCCGTGC AGTCCGCTCT TCACTGTTCC ACGGCCTCCC AGTGCCTCCC AGCATTGGAC 1320
CCATCTCCCC CTGCAGTTTG AGGCCAGAGA GGTGAGTGGA CCTGACAAGT GCCAGAGTAA 1380
CCGTGTAGAC AGAGCAGTGT AGACAGCGCT CAGCCCCAGC CCCAGGTGTG GACCTCATGC 1440
TGGTGATGGC TCCCCTGGGT GGCCTGCCAG CACAGCCAGT KCCATCAGGG AGCTGAAGGG 1500
GCTGTCCCCC ACCTAACTCC AGCTCCCCCT TCACGTTGTC ACCAAGGCCC TGTGCCGCCC 1560
GCCTCGCCCC CCTGCTCTGT GGATTCCTTT GGAAGGGCT CCCTGGGCAG GACAATAAAG 1620
AGT'TTTGACT CCAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1670

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 648 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

CCGACTCCAG CTCTGAGCCT GTTCGCGGCT CTCGGCTTCC ACTGCAGCCA TGCTACTCCT      60
CTTGCTGGTG GTCTCAGCCC TTCACATCCT CATTCTTATA CTGCTTTTCG TGGCCACTTT      120
GGACAAGTCC TGGTGGACTC TCCCTGGGAA AGAGTCCCTG AATCTCTGGT ACGACTGCAC      180
GTGGAACAAC GACACCAAAA CATGGGCCTG CAGTAATGTC AGCGAGAATG GCTGGCTGAA      240
GGCGGTGCAG GTCCTCATGG TGCTCTCCCT CATTCTCTGC TGTCTCTCCT TCATCCTGTT      300
CATGTTCCAG CTCTACACCA TGCGACGAGG AGGTCTCTTC TATGCCACCG GCCTCTGCCA      360
GCTTTGCACC AGCGTGGCGG TGTTTACTGG CGCCTTGATC TATGCCATTC ACGCCGAGGA      420
GATCCTGGAG AAGCACCCGC GAGGGGGCAG CTCGGGATAC TGCTTCGCCC TGGCCTGGGT      480
GGCCTTCCCC CTCGCCCTGG TCAGCGGCAT CATCTACATC CACCTACGGA AGCGGGAGTG      540
AGCGCCCCGC CTCGCTCGGC TGCCCCCGCC CCTTCCCGGC CCCCCTCGCC GCGCGTCCTC      600
CAAAAAATAA AACCTTAACC GCGAAAAAAA AAAAAAAAAA AAAAAAAA      648

```

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 163 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

Met Ser Leu Leu Leu Val Val Ser Ala Leu His Ile Leu Ile Leu
1           5           10           15
Ile Leu Leu Phe Val Ala Thr Leu Asp Lys Ser Trp Trp Thr Leu Pro
20           25           30
Gly Lys Glu Ser Leu Asn Leu Trp Tyr Asp Cys Thr Trp Asn Asn Asp
35           40           45
Thr Lys Thr Trp Ala Cys Ser Asn Val Ser Glu Asn Gly Trp Leu Lys
50           55           60

```

Ala Val Gln Val Leu Met Val Leu Ser Leu Ile Leu Cys Cys Leu Ser
65 70 75 80

Phe Ile Leu Phe Met Phe Gln Leu Tyr Thr Met Arg Arg Gly Gly Leu
85 90 95

Phe Tyr Ala Thr Gly Leu Cys Gln Leu Cys Thr Ser Val Ala Val Phe
100 105 110

Thr Gly Ala Leu Ile Tyr Ala Ile His Ala Glu Glu Ile Leu Glu Lys
115 120 125

His Pro Arg Gly Gly Ser Phe Gly Tyr Cys Phe Ala Leu Ala Trp Val
130 135 140

Ala Phe Pro Leu Ala Leu Val Ser Gly Ile Ile Tyr Ile His Leu Arg
145 150 155 160

Lys Arg Glu

What is claimed is:

1. A composition comprising an isolated polynucleotide selected from the group consisting of:
 - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
 - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 370 to nucleotide 1341;
 - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 700 to nucleotide 1341;
 - (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 370 to nucleotide 798;
 - (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone B121 deposited under accession number ATCC 98019;
 - (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone B121 deposited under accession number ATCC 98019;
 - (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone B121 deposited under accession number ATCC 98019;
 - (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone B121 deposited under accession number ATCC 98019;
 - (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
 - (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
 - (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
 - (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
 - (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).
2. A composition of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with a composition of claim 2.
4. The host cell of claim 3, wherein said cell is a mammalian cell.
5. A process for producing a protein encoded by a composition of claim 2, which process comprises:
 - (a) growing a culture of the host cell of claim 3 in a suitable culture medium; and
 - (b) purifying said protein from the culture.
6. A protein produced according to the process of claim 5.
7. The protein of claim 6 comprising a mature protein.
8. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:
 - (a) the amino acid sequence of SEQ ID NO:2;
 - (b) the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 143;
 - (c) fragments of the amino acid sequence of SEQ ID NO:2; and
 - (d) the amino acid sequence encoded by the cDNA insert of clone B121 deposited under accession number ATCC 98019;the protein being substantially free from other mammalian proteins.
9. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.
10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 1 to amino acid 143.
11. The composition of claim 8, further comprising a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

14. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 24 to nucleotide 548;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 525 to nucleotide 548;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 115 to nucleotide 317;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone B196 deposited under accession number ATCC 98021;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone B196 deposited under accession number ATCC 98021;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone B196 deposited under accession number ATCC 98021;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone B196 deposited under accession number ATCC 98021;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
- (b) the amino acid sequence of SEQ ID NO:4 from amino acid 118 to amino acid 162;
- (c) fragments of the amino acid sequence of SEQ ID NO:4; and
- (d) the amino acid sequence encoded by the cDNA insert of clone B196 deposited under accession number ATCC 98021;

the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3 or SEQ ID NO:5.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 50 to nucleotide 538;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 290 to nucleotide 538;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone D157 deposited under accession number ATCC 98020;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone D157 deposited under accession number ATCC 98020;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone D157 deposited under accession number ATCC 98020;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone D157 deposited under accession number ATCC 98020;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:7 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

18. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:7;

(b) fragments of the amino acid sequence of SEQ ID NO:7; and

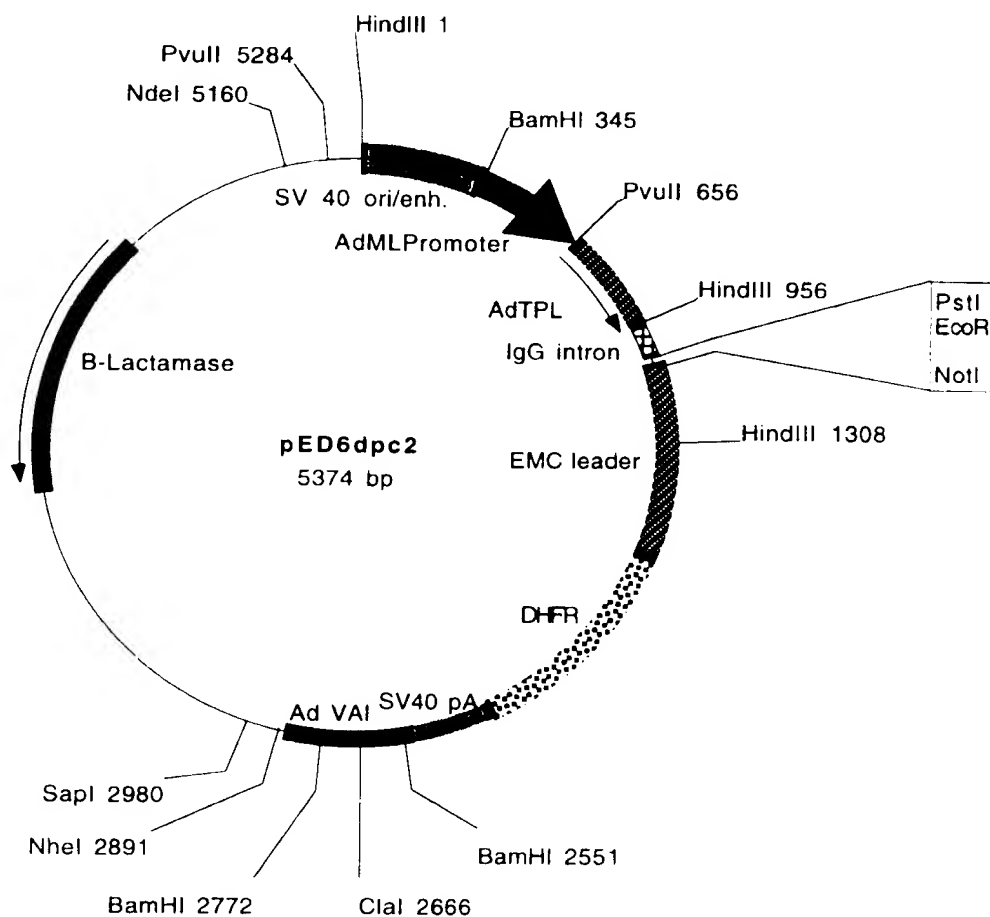
(c) the amino acid sequence encoded by the cDNA insert of clone D157

deposited under accession number ATCC 98020;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:6.

FIGURE 1A

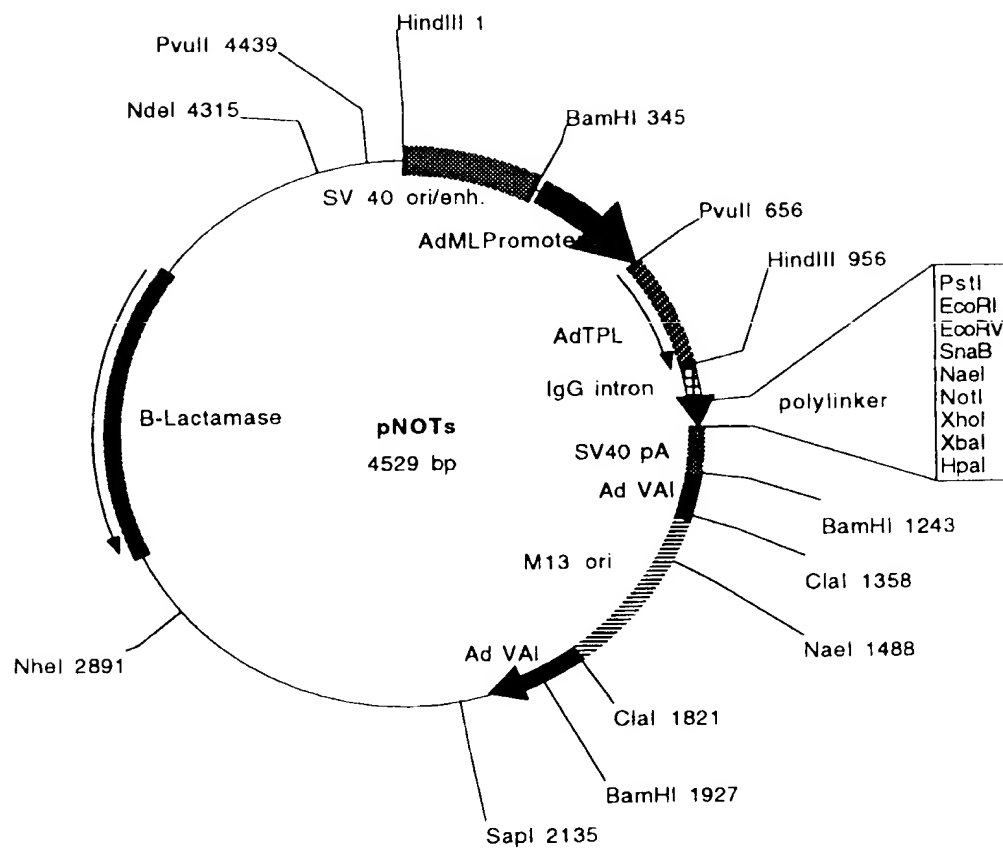


Plasmid name: pED6dpc2

Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

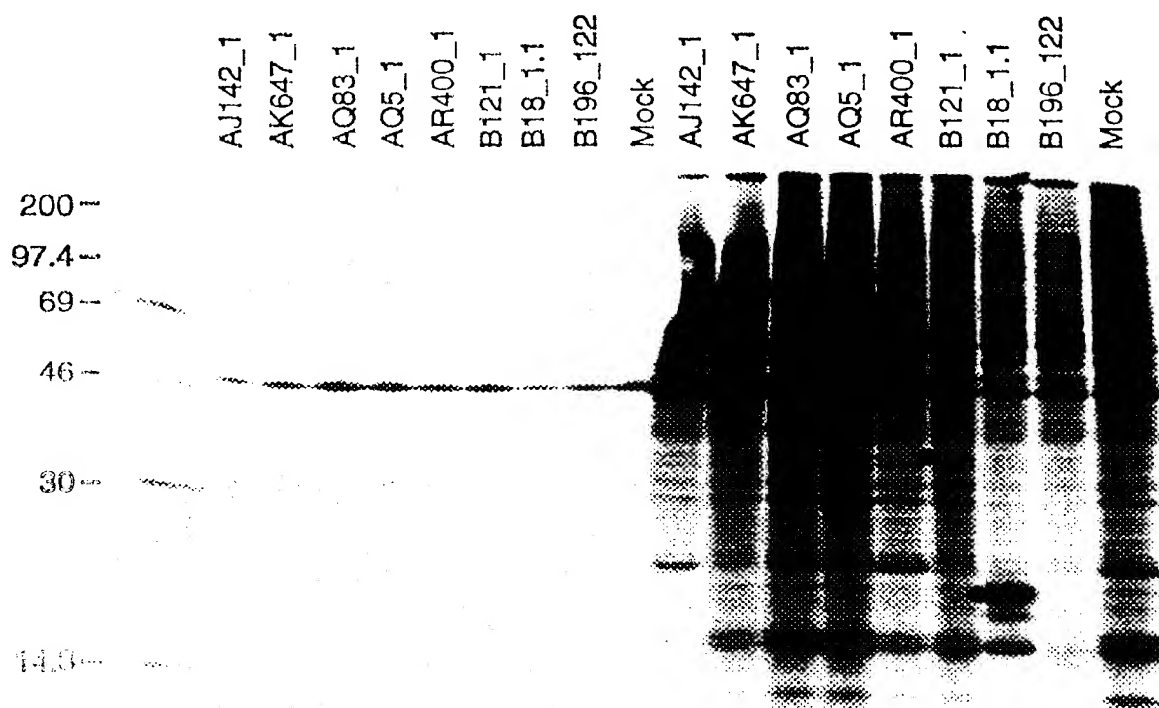
FIGURE 1B



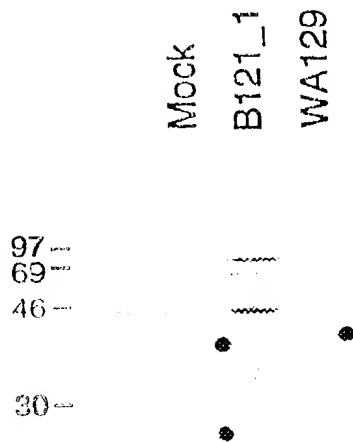
Plasmid name: pNOTs

Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al, 1989. Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the Clal site. SST cDNAs are cloned between EcoRI and NotI



3/4
Fig. 2



4/4

Fig. 3

SUBSTITUTE SHEET (RULE 26)



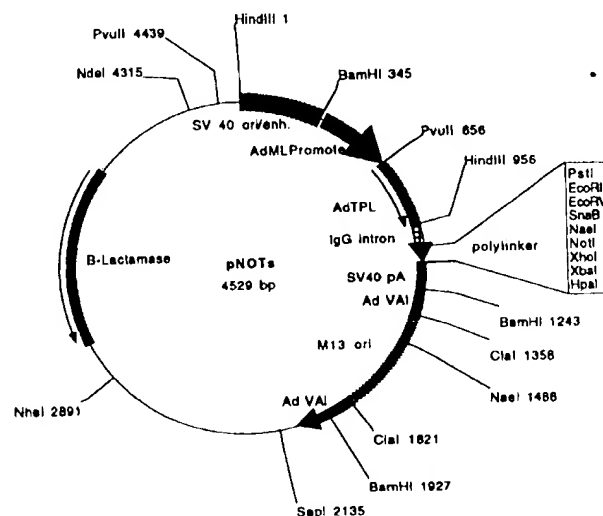
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 5/10, C07K 14/47, C12Q 1/68, A61K 38/17		A3		(11) International Publication Number: WO 98/30696
				(43) International Publication Date: 16 July 1998 (16.07.98)
(21) International Application Number: PCT/US98/00575		(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).		
(22) International Filing Date: 13 January 1998 (13.01.98)				
(30) Priority Data: 08/783,395 13 January 1997 (13.01.97) AP (34) Countries for which the regional or international application was filed: US et al. PCT/US97/05682 4 April 1997 (04.04.97) WO (34) Countries for which the regional or international application was filed: JP et al. 08/866,022 30 May 1997 (30.05.97) US 08/924,838 5 September 1997 (05.09.97) US 09/005,986 12 January 1998 (12.01.98) US		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).		
(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.		
(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).		(88) Date of publication of the international search report: 29 October 1998 (29.10.98)		

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al, 1989, Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SSI cDNAs are cloned between EcoRI and NotI.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

 Inter-Application No
 PC 1 98/00575

 A. CLASSIFICATION OF SUBJECT MATTER
 IPC 6 C12N15/12 C12N5/10 C07K14/47 C12Q1/68 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

 Minimum documentation searched (classification system followed by classification symbols)
 IPC 6 C12N C07K C12Q A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	L. HILLIER ET AL.: "zd68f10.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 345835 5'" EMBL SEQUENCE DATABASE, 25 June 1996, HEIDELBERG, FRG, XP002060222 Accession no. W77809	1,13
X	C. AUFRAY ET AL.: "H. sapiens partial cDNA sequence; clone c-0ja07" EMBL SEQUENCE DATABASE, 5 November 1994, HEIDELBERG, FRG, XP002060223 Accession no. Z42428	1,13

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

2 June 1998

Date of mailing of the international search report

11.09.1998

Name and mailing address of the ISA

 European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

HORNIG H.

INTERNATIONAL SEARCH REPORT

International Application No

PC, /US 98/00575

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X, L	WO 97 38084 A (GENETICS INST) 16 October 1997 SEQ ID nos. 1 and 2 see claims 1-12 ---	1-13
E	WO 98 04693 A (GENETICS INST ;JACOBS KENNETH (US)) 5 February 1998 SEQ ID nos. 23 and 24 see claims 20-22 ---	1-13
A	ADAMS M D ET AL: "3,400 NEW EXPRESSED SEQUENCE TAGS IDENTIFY DIVERSITY OF TRANSCRIPTS IN HUMAN BRAIN" NATURE GENETICS, vol. 4, no. 3, pages 256-267, XP000611495 see the whole document ---	1-13
A	JACOBS K ET AL: "A novel method for isolating eukaryotic cDNA clones encoding secreted proteins." KEYSTONE SYMPOSIUM ON DENDRITIC CELLS: ANTIGEN PRESENTING CELLS OF T AND B LYMPHOCYTES, TAOS, NEW MEXICO, USA, MARCH 10-16, 1995. JOURNAL OF CELLULAR BIOCHEMISTRY SUPPLEMENT 0 (21A). 1995. 19. ISSN: 0733-1959, XP002027246 abstract no. C1-207 see abstract ---	1-13
A	EP 0 510 691 A (OSAKA BIOSCIENCE INST) 28 October 1992 see the whole document ---	1-13
A	WO 94 07916 A (MERCK & CO INC ;SCHMIDT AZRIEL (US); RODAN GIDEON A (US); RUTLEDGE) 14 April 1994 see the whole document ---	1-13
A	WO 90 05780 A (OREGON STATE) 31 May 1990 see the whole document ---	1-13
A	WO 90 14432 A (GENETICS INST) 29 November 1990 see the whole document ---	1-13
A	WO 96 17925 A (IMMUNEX CORP) 13 June 1996 see the whole document ---	1-13

-/--

INTERNATIONAL SEARCH REPORT

Inter. Patent Application No.

PCT/98/00575

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	R.J. KAUFMAN ET AL.: "Effect of von Willebrand factor coexpression on the synthesis and secretion of factor VIII in chinese hamster ovary cells" MOL. CELL. BIOL., vol. 9, no. 3, March 1989, ASM WASHINGTON, DC,US, pages 1233-1242, XP002041592 see the whole document ---	1-13
A	R.J. KAUFMAN ET AL.: "The phosphorylation state of eucaryotic initiation factor 2 alters translation efficiency of specific mRNAs" MOL. CELL. BIOL., vol. 9, no. 3, March 1989, ASM WASHINGTON, DC,US, pages 946-958, XP002041593 see the whole document ---	1-13
A	R.J. KAUFMAN ET AL.: "Improved vectors for stable expression of foreign genes in mammalian cells by use of the untranslated leader sequence from EMC virus" NUCLEIC ACIDS RESEARCH, vol. 19, no. 16, 1991, IRL PRESS LIMITED,OXFORD,ENGLAND, pages 4485-4490, XP002041594 cited in the application see the whole document ---	1-13
A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 see the whole document ---	1-13
P,A	WO 97 07198 A (GENETICS INSTITUT) 27 February 1997 see the whole document -----	1-13

INTERNATIONAL SEARCH REPORT

national application No.

PCT/US 98/ 00575

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim 12 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see continuation sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-13

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-13

A composition comprising an isolated polynucleotide selected from the group consisting of: SEQ ID no.1; said composition wherein said polynucleotide is operably linked to an expression control sequence; a host cell transformed with said composition; a process for producing a protein which is encoded by said polynucleotide sequence; a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group of SEQ ID no.2, said composition further comprising a pharmaceutical acceptable carrier; a method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of said composition, the gene corresponding to the cDNA sequence of SEQ ID no.1.

2. Claims: 14-16

A composition comprising an isolated polynucleotide sequence selected from the group of SEQ ID no.3; a composition comprises a protein, wherein said protein comprises an amino acid sequence selected from the group of SEQ ID no.4; the gene corresponding to the cDNA sequences of SEQ ID no.3 or SEQ ID no.5;

3. Claims: 17-19

Idem as subject 2 but limited to SEQ ID nos.6 and 7.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/00575

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9738084 A	16-10-97	AU 2723397 A	29-10-97
WO 9804693 A	05-02-98	US 5708157 A	13-01-98
		AU 3890797 A	20-02-98
		AU 3890897 A	20-02-98
		AU 3962497 A	20-02-98
		AU 4044297 A	20-02-98
		WO 9804694 A	05-02-98
		WO 9804695 A	05-02-98
		WO 9804696 A	05-02-98
		US 5786465 A	28-07-98
EP 0510691 A	28-10-92	CA 2067031 A	27-10-92
		JP 5184368 A	27-07-93
WO 9407916 A	14-04-94	AU 5165193 A	26-04-94
WO 9005780 A	31-05-90	AT 154636 T	15-07-97
		AU 645963 B	03-02-94
		AU 4668689 A	12-06-90
		DE 68928137 D	24-07-97
		DE 68928137 T	19-02-98
		EP 0447483 A	25-09-91
		ES 2104599 T	16-10-97
		JP 4506449 T	12-11-92
WO 9014432 A	29-11-90	US 5580753 A	03-12-96
		AT 147436 T	15-01-97
		AU 637620 B	03-06-93
		AU 5928990 A	18-12-90
		CA 2056997 A	24-11-90
		DE 69029657 D	20-02-97
		DK 473724 T	14-04-97
		EP 0473724 A	11-03-92
		ES 2099096 T	16-05-97
		JP 4506006 T	22-10-92
		US 5734037 A	31-03-98
		US 5414071 A	09-05-95
WO 9617925 A	13-06-96	AU 4639396 A	26-06-96

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte Application No

PCT/98/00575

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9617925 A		CA 2206488 A FI 972390 A NO 972455 A	13-06-96 05-06-97 06-08-97
US 5536637 A	16-07-96	US 5712116 A	27-01-98
WO 9707198 A	27-02-97	US 5707829 A AU 6712396 A AU 6768596 A EP 0839196 A EP 0851875 A WO 9704097 A	13-01-98 18-02-97 12-03-97 06-05-98 08-07-98 06-02-97



...

...

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/00, C07K 14/47, G01N 33/53		A2	(11) International Publication Number: WO 00/52151
			(43) International Publication Date: 8 September 2000 (08.09.00)
(21) International Application Number: PCT/US00/05621		Belmont Place, San Jose, CA 95136 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US).	
(22) International Filing Date: 3 March 2000 (03.03.00)		(74) Agents: HAMLET-COX, Diana et al.; Incyte Pharmaceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).	
(30) Priority Data: 60/123,117 5 March 1999 (05.03.99) US		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 60/123,117 (CIP) Filed on 5 March 1999 (05.03.99)		Published Without international search report and to be republished upon receipt of that report.	
(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). LU, Dyung, Aina, M. [US/US]; 55 Park			
(54) Title: HUMAN SECRETORY PROTEINS			
(57) Abstract			
<p>The invention provides human secretory proteins (HSECP) and polynucleotides which identify and encode HSECP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HSECP.</p>			



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		



HUMAN SECRETORY PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human secretory proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders.

BACKGROUND OF THE INVENTION

Protein transport and secretion are essential for cellular function. Protein transport is mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Secreted proteins are often synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include receptors, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, neuropeptides, vasomediators, ion channels, transporters/pumps, and proteases. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

G-protein coupled receptors (GPCRs) comprise a superfamily of integral membrane proteins which transduce extracellular signals. Not all GPCRs contain N-terminal signal peptides. GPCRs include receptors for biogenic amines such as dopamine, epinephrine, histamine, glutamate (metabotropic-type), acetylcholine (muscarinic-type), and serotonin; for lipid mediators of inflammation such as prostaglandins, platelet activating factor, and leukotrienes; for peptide hormones such as calcitonin, C5a anaphylatoxin, follicle stimulating hormone, gonadotropin releasing hormone, neurokinin, oxytocin, and thrombin; and for sensory signal mediators such as retinal photopigments and olfactory stimulatory molecules. The structure of these highly conserved receptors consists of seven hydrophobic transmembrane regions, cysteine disulfide bridges between the second and third extracellular loops, an extracellular N-terminus, and a cytoplasmic C-terminus.



The N-terminus interacts with ligands, the disulfide bridges interact with agonists and antagonists, and the large third intracellular loop interacts with G proteins to activate second messengers such as cyclic AMP, phospholipase C, inositol triphosphate, or ion channels. (Reviewed in Watson, S. and Arkinstall, S. (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego, CA, pp. 2-6; and Bolander, F.F. (1994) Molecular Endocrinology, Academic Press, San Diego, CA, pp. 162-176.)

Other types of receptors include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, MPs may be heavily glycosylated and may contain an Arginine-Glycine-Aspartate (RGD) tripeptide motif which may play a role in adhesive interactions. MPs include extracellular proteins such as fibronectin, collagen, galectin, vitronectin and its proteolytic derivative somatomedin B; and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) *Kidney Int.* 51:1413-1417; Sjaastad, M.D. and Nelson, W.J. (1997) *BioEssays* 19:47-55.)

Cytokines are secreted by hematopoietic cells in response to injury or infection. Interleukins, neurotrophins, growth factors, interferons, and chemokines all define cytokine families that work in



conjunction with cellular receptors to regulate cell proliferation and differentiation. In addition, cytokines effect activities such as leukocyte migration and function, hematopoietic cell proliferation, temperature regulation, acute response to infection, tissue remodeling, and apoptosis.

Chemokines, in particular, are small chemoattractant cytokines involved in inflammation, leukocyte proliferation and migration, angiogenesis and angiostasis, regulation of hematopoiesis, HIV infectivity, and stimulation of cytokine secretion. Chemokines generally contain 70-100 amino acids and are subdivided into four subfamilies based on the presence of conserved cysteine-based motifs. (Callard, R. and Gearing, A. (1994) The Cytokine Facts Book, Academic Press, New York, NY, pp. 181-190, 210-213, 223-227.)

Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with MPs for activity. Complex interactions among these factors and their receptors trigger intracellular signal transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth factors such as epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, and platelet-derived growth factor. The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs). Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils, neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin which function as hormones to regulate cellular functions other than proliferation.

Growth and differentiation factors play critical roles in neoplastic transformation of cells in vitro and in tumor progression in vivo. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both in vivo and in vitro. Moreover, some growth factors and growth factor receptors are related both structurally and functionally to oncoproteins. In addition, growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)

Proteolytic enzymes or proteases either activate or deactivate proteins by hydrolyzing peptide bonds. Proteases are found in the cytosol, in membrane-bound compartments, and in the extracellular space. The major families are the zinc, serine, cysteine, thiol, and carboxyl proteases.

Ion channels, ion pumps, and transport proteins mediate the transport of molecules across



cellular membranes. Transport can occur by a passive, concentration-dependent mechanism or can be linked to an energy source such as ATP hydrolysis. Symporters and antiporters transport ions and small molecules such as amino acids, glucose, and drugs. Symporters transport molecules and ions unidirectionally, and antiporters transport molecules and ions bidirectionally. Transporter
5 superfamilies include facilitative transporters and active ATP-binding cassette transporters which are involved in multiple-drug resistance and the targeting of antigenic peptides to MHC Class I molecules. These transporters bind to a specific ion or other molecule and undergo a conformational change in order to transfer the ion or molecule across the membrane. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 523-546.)

10 Ion channels are formed by transmembrane proteins which create a lined passageway across the membrane through which water and ions, such as Na^+ , K^+ , Ca^{2+} , and Cl^- , enter and exit the cell. For example, chloride channels are involved in the regulation of the membrane electric potential as well as absorption and secretion of ions across the membrane. Chloride channels also regulate the internal pH of membrane-bound organelles.

15 Ion pumps are ATPases which actively maintain membrane gradients. Ion pumps are classified as P, V, or F according to their structure and function. All have one or more binding sites for ATP in their cytosolic domains. The P-class ion pumps include Ca^{2+} ATPase and Na^+/K^+ ATPase and function in transporting H^+ , Na^+ , K^+ , and Ca^{2+} ions. P-class pumps consist of two α and two β transmembrane subunits. The V- and F-class ion pumps have similar structures but transport only H^+ .
20 F class H^+ pumps mediate transport across the membranes of mitochondria and chloroplasts, while V-class H^+ pumps regulate acidity inside lysosomes, endosomes, and plant vacuoles.

A family of structurally related intrinsic membrane proteins known as facilitative glucose transporters catalyze the movement of glucose and other selected sugars across the plasma membrane. The proteins in this family contain a highly conserved, large transmembrane domain comprised of 12
25 α -helices, and several weakly conserved, cytoplasmic and exoplasmic domains. (Pessin, J. E., and Bell, G.I. (1992) *Annu. Rev. Physiol.* 54:911-930.)

Amino acid transport is mediated by Na^+ dependent amino acid transporters. These transporters are involved in gastrointestinal and renal uptake of dietary and cellular amino acids and in neuronal reuptake of neurotransmitters. Transport of cationic amino acids is mediated by the
30 system y^+ family and the cationic amino acid transporter (CAT) family. Members of the CAT family share a high degree of sequence homology, and each contains 12-14 putative transmembrane domains. (Ito, K. and Groudine, M. (1997) *J. Biol. Chem.* 272:26780-26786.)

Hormones are secreted molecules that travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical
35 compositions and mechanisms of action, hormones can be grouped into two categories. One category



includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes

- 5 hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotrophic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific
10 American Books Inc., New York, NY, pp. 856-864.)

- Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin,
15 vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C. R. et al. (1985) Endocrine
20 Physiology, Oxford University Press, New York, NY, pp. 57-62.)

The discovery of new human secretory proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders.

25

SUMMARY OF THE INVENTION

- The invention features purified polypeptides, human secretory proteins, referred to collectively as "HSECP" and individually as "HSECP-1," "HSECP-2," "HSECP-3," "HSECP-4," "HSECP-5," "HSECP-6," "HSECP-7," "HSECP-8," "HSECP-9," "HSECP-10," "HSECP-11,"
30 "HSECP-12," "HSECP-13," "HSECP-14," "HSECP-15," "HSECP-16," "HSECP-17," "HSECP-18," "HSECP-19," "HSECP-20," "HSECP-21" and "HSECP-22." In one aspect, the invention provides an isolated polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active
35 fragment of an amino acid sequence selected from the group consisting of SEQ ID NO.1-22, or d) an



immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-22.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising

- 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. In one
10 alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:23-44.

- Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group
15 consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

- 20 The invention also provides a method for producing a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino
25 acid sequence selected from the group consisting of SEQ ID NO:1-22. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

- Additionally, the invention provides an isolated antibody which specifically binds to a
30 polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID
35 NO:1-22.



The invention further provides an isolated polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44, c) a polynucleotide sequence complementary to a), or d) a polynucleotide sequence complementary to b). The method comprises a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 30 contiguous nucleotides. In another alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, and a pharmaceutically acceptable excipient. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional HSECP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. The method comprises a) exposing a sample comprising the polypeptide to a



compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional HSECP.

5 comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, 10 c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, or d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by 15 the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional HSECP, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention further provides a method for screening a compound for effectiveness in 20 altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:23-44, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

25 BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HSECP.

Table 2 shows features of each polypeptide sequence, including predicted signal peptides and 30 other motifs, and methods, algorithms, and searchable databases used for analysis of HSECP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

35 Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones



encoding HSECP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze HSECP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"HSECP" refers to the amino acid sequences of substantially purified HSECP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of HSECP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HSECP either by directly interacting with HSECP or by acting on components of the biological pathway in which HSECP participates.

An "allelic variant" is an alternative form of the gene encoding HSECP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to



allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

“Altered” nucleic acid sequences encoding HSECP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as HSECP or a polypeptide with at least one functional characteristic of HSECP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HSECP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HSECP. The encoded protein may also be “altered,” and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HSECP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HSECP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms “amino acid” and “amino acid sequence” refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where “amino acid sequence” is recited to refer to an amino acid sequence of a naturally occurring protein molecule, “amino acid sequence” and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

“Amplification” relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term “antagonist” refers to a molecule which inhibits or attenuates the biological activity of HSECP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of HSECP either by directly interacting with HSECP or by acting on components of the biological pathway in which HSECP participates.

The term “antibody” refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant.



Antibodies that bind HSECP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired.

5 Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to
10 immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense"
15 strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be
20 produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

25 The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HSECP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of
30 polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization
35 between the nucleic acid strands. This is of particular importance in amplification reactions, which



depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HSECP or fragments of HSECP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
25	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
30	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
35	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
40	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr



Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

5 A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative
10 polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of HSECP or the polynucleotide encoding HSECP which is
15 identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid
20 residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present
25 embodiments.

A fragment of SEQ ID NO:23-44 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:23-44, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:23-44 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:23-44 from related
30 polynucleotide sequences. The precise length of a fragment of SEQ ID NO:23-44 and the region of SEQ ID NO:23-44 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-22 is encoded by a fragment of SEQ ID NO:23-44. A fragment of SEQ ID NO:1-22 comprises a region of unique amino acid sequence that specifically identifies
35 SEQ ID NO:1-22. For example, a fragment of SEQ ID NO:1-22 is useful as an immunogenic peptide



for the development of antibodies that specifically recognize SEQ ID NO:1-22. The precise length of a fragment of SEQ ID NO:1-22 and the region of SEQ ID NO:1-22 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

5 The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization
10 assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to
15 one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

20 The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

25 Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS
30 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

 Alternatively, a suite of commonly used and freely available sequence comparison algorithms
35 is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment



Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the



site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the



stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%.

Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_{ot} or R_{ot} analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune



disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of HSECP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of HSECP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of HSECP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HSECP.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Probe" refers to nucleic acid sequences encoding HSECP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes.

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target



DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to



identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence.

- 5 This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter
- 10 sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

- 15 An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

- The term "sample" is used in its broadest sense. A sample suspected of containing nucleic
- 20 acids encoding HSECP, or fragments thereof, or HSECP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

- The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or
- 25 synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

- 30 The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

- A "substitution" refers to the replacement of one or more amino acids or nucleotides by
- 35 different amino acids or nucleotides, respectively.



"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

5 "Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral
10 infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to
15 animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in
20 vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, and plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention
25 into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-
30 1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of
35 polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding



polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human secretory proteins (HSECP), the polynucleotides encoding HSECP, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HSECP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HSECP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each HSECP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; and column 6 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 6 were used to characterize each polypeptide through sequence homology and protein motifs. In column 5, the first line of each cell lists the amino acid residues comprising predicted signal peptide sequences located at the amino terminus of each



HSECP. Additional identifying motifs or signatures, such as a somatomedin B signature in SEQ ID NO:16 and seven putative transmembrane domains in SEQ ID NO:18, are also listed in column 5.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HSECP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:23-44 and to distinguish between SEQ ID NO:23-44 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express HSECP as a fraction of total tissues expressing HSECP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing HSECP as a fraction of total tissues expressing HSECP. Column 5 lists the vectors used to subclone each cDNA library. In particular, three out of four cDNA libraries which express SEQ ID NO:23 are derived from cartilage and synovia associated with joint inflammation, and four out of five cDNA libraries which express SEQ ID NO:29 are derived from intestinal tissue. Furthermore, about half of the cDNA libraries expressing SEQ ID NO:34 are associated with inflammation or the hematopoietic/immune system. Likewise, about half of the cDNA libraries expressing SEQ ID NO:35 are associated with inflammation or the hematopoietic/immune system, and in particular, with inflammation of the joints. In addition, 82% of the cDNA libraries expressing SEQ ID NO:37 are derived from tissues of the nervous system. Finally, expression of SEQ ID NO:39 is detected solely in a subtracted prostate tumor cDNA library, and expression of SEQ ID NO:43 is detected only in two cDNA libraries derived from heart tissue.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding HSECP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses HSECP variants. A preferred HSECP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the HSECP amino acid sequence, and which contains at least one functional or structural characteristic of HSECP.

The invention also encompasses polynucleotides which encode HSECP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:23-44, which encodes HSECP. The polynucleotide sequences of SEQ ID NO:23-44, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.



The invention also encompasses a variant of a polynucleotide sequence encoding HSECP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HSECP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:23-44 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:23-44. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HSECP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HSECP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HSECP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HSECP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring HSECP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HSECP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HSECP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HSECP and HSECP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HSECP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:23-44 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and



G33S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

5 Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is
10 automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a
15 variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HSECP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences,
20 such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising
25 a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region
30 of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed
35 using commercially available software, such as OLIGO 4.06 Primer Analysis software (National



Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HSECP may be cloned in recombinant DNA molecules that direct expression of HSECP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HSECP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HSECP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of HSECP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene



variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding HSECP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser. 7*:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser. 7*:225-232.) Alternatively, HSECP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HSECP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active HSECP, the nucleotide sequences encoding HSECP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HSECP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HSECP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HSECP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding



sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used.

5 (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HSECP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A

10 Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HSECP. These include, but are not limited to, microorganisms such as bacteria transformed
15 with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

20 In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HSECP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HSECP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPT11 plasmid (Life Technologies). Ligation of sequences encoding HSECP into the vector's multiple
25 cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of HSECP are needed, e.g. for the production of
30 antibodies, vectors which direct high level expression of HSECP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HSECP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such
35 vectors direct either the secretion or intracellular retention of expressed proteins and enable



integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

- Plant systems may also be used for expression of HSECP. Transcription of sequences encoding HSECP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.)
- 10 These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

- In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HSECP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HSECP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.
- 20

- Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)
- 25

- For long term production of recombinant proteins in mammalian systems, stable expression of HSECP in cell lines is preferred. For example, sequences encoding HSECP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector.
- 30 Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

- 35 Any number of selection systems may be used to recover transformed cell lines. These



include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apv* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HSECP is inserted within a marker gene sequence, transformed cells containing sequences encoding HSECP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HSECP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HSECP and that express HSECP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HSECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HSECP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana



Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HSECP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HSECP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HSECP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HSECP may be designed to contain signal sequences which direct secretion of HSECP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HSECP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HSECP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HSECP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST).



maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HSECP encoding sequence and the heterologous protein sequence, so that HSECP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10).
10 A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HSECP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.
15

Fragments of HSECP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of HSECP may be synthesized separately and then combined to produce the full length molecule.
20

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HSECP and human secretory proteins. In addition, the expression of HSECP is closely associated with cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders. Therefore, HSECP appears to play a role in cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders. In the treatment of disorders associated with increased HSECP expression or activity, it is desirable to decrease the expression or activity of HSECP. In the treatment of disorders associated with decreased HSECP expression or activity, it is desirable to increase the expression or activity of HSECP.
25
30

Therefore, in one embodiment, HSECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSECP. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall
35



- bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma,
- 5 atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's
- 10 thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal
- 15 circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis,
- 20 pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic
- 25 encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a
- 30 cardiovascular disorder, and in particular, a disorder of the heart such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus
- 35 erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart



disease, congenital heart disease, and complications of cardiac transplantation; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder.

In another embodiment, a vector capable of expressing HSECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSECP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HSECP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSECP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HSECP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSECP including, but not limited to, those listed above.

In a further embodiment, an antagonist of HSECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HSECP. Examples of such disorders include, but are not limited to, those cancer, inflammation, and gastrointestinal, cardiovascular, and neurological disorders described above. In one aspect, an antibody which specifically binds HSECP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express HSECP.

In an additional embodiment, a vector expressing the complement of the polynucleotide



encoding HSECP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HSECP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HSECP may be produced using methods which are generally known in the art. In particular, purified HSECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HSECP. Antibodies to HSECP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HSECP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HSECP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HSECP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HSECP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J.

Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and



Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HSECP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g.,

10 Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

15 Antibody fragments which contain specific binding sites for HSECP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

20 Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HSECP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HSECP epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HSECP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of HSECP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HSECP epitopes, represents the average affinity, or avidity, of the antibodies for HSECP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HSECP epitope, represents a true measure of affinity. High-affinity antibody

30
35



preparations with K_d ranging from about 10^6 to 10^{12} L/mole are preferred for use in immunoassays in which the HSECP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_d ranging from about 10^8 to 10^{10} L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HSECP.

- 5 preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For
10 example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of HSECP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

- 15 In another embodiment of the invention, the polynucleotides encoding HSECP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HSECP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HSECP. Thus, complementary molecules or
20 fragments may be used to modulate HSECP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HSECP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or
25 from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HSECP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

- Genes encoding HSECP can be turned off by transforming a cell or tissue with expression
30 vectors which express high levels of a polynucleotide, or fragment thereof, encoding HSECP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of
35 the vector system.



As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5' or regulatory regions of the gene encoding HSECP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed.

- 5 Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco
10 NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

- Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example,
15 engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HSECP.

- Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides,
20 corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

- Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared
25 by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HSECP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA
30 constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

- RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2'O-methyl rather than phosphodiesterase
35 linkages within the backbone of the molecule. This concept is inherent in the production of PNAs



and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

5 Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat.

10 Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

15 An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HSECP, antibodies to HSECP, and mimetics, agonists, antagonists, or inhibitors of HSECP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, 20 but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, 25 enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's 30 Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

35 Pharmaceutical preparations for oral use can be obtained through combining active



compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many



acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HSECP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HSECP or fragments thereof, antibodies of HSECP, and agonists, antagonists or inhibitors of HSECP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular



formulation.

Normal dosage amounts may vary from about 0.1 μ g to 100,000 μ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.

- 5 Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

- In another embodiment, antibodies which specifically bind HSECP may be used for the
- 10 diagnosis of disorders characterized by expression of HSECP, or in assays to monitor patients being treated with HSECP or agonists, antagonists, or inhibitors of HSECP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HSECP include methods which utilize the antibody and a label to detect HSECP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or
- 15 without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HSECP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HSECP expression.

- 20 Normal or standard values for HSECP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to HSECP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of HSECP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values.
- 25 Deviation between standard and subject values establishes the parameters for diagnosing disease.

- In another embodiment of the invention, the polynucleotides encoding HSECP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of HSECP may be correlated
- 30 with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HSECP, and to monitor regulation of HSECP levels during therapeutic intervention.

- In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HSECP or closely related molecules may be used to identify nucleic acid sequences which encode HSECP. The specificity of the probe, whether it is
- 35 made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a



conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding HSECP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the HSECP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:23-44 or from genomic sequences including promoters, enhancers, and introns of the HSECP gene.

Means for producing specific hybridization probes for DNAs encoding HSECP include the cloning of polynucleotide sequences encoding HSECP or HSECP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HSECP may be used for the diagnosis of disorders associated with expression of HSECP. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea,



- emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis,
- 5 Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein
- 10 obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a cardiovascular disorder, and in particular, a disorder of the heart such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease,
- 15 degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation; and a neurological
- 20 disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess,
- 25 suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central
- 30 nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia,
- 35 catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic



neuralgia, and Tourette's disorder. The polynucleotide sequences encoding HSECP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HSECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HSECP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HSECP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HSECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HSECP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HSECP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.



Additional diagnostic uses for oligonucleotides designed from the sequences encoding HSECP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HSECP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HSECP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of HSECP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HSECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the



Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding HSECP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HSECP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HSECP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HSECP, or fragments thereof, and washed. Bound HSECP is then detected by methods well known in the art. Purified HSECP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HSECP specifically compete with a test compound for binding HSECP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HSECP.

In additional embodiments, the nucleotide sequences which encode HSECP may be used in



any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific
5 embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific
10 embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/123,117, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a
20 monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A⁺) RNA was isolated
25 using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA
30 libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the
35 appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-



1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.



The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:23-44. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs



from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HSECP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of HSECP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:23-44 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR



was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and β -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM



BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:23-44 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

5 VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:23-44 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06
10 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based
15 hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature
20 under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array
25 elements on the surface of a substrate. (See, e.g., Baldeschweiler, *supra*.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and
30 patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software
35 well known in the art such as LASERGENE software (DNASTAR). Full length cDNAs, ESTs, or



fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the HSECP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HSECP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HSECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HSECP-encoding transcript.

IX. Expression of HSECP

Expression and purification of HSECP is achieved using bacterial or virus-based expression systems. For expression of HSECP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3).

Antibiotic resistant bacteria express HSECP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HSECP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HSECP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)



In most expression systems, HSECP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HSECP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified HSECP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of HSECP Activity

An assay for HSECP activity measures the expression of HSECP on the cell surface. cDNA encoding HSECP is subcloned into an appropriate mammalian expression vector suitable for high levels of cDNA expression. The resulting construct is transfected into a nonhuman cell line such as NIH3T3. Cell surface proteins are labeled with biotin using methods known in the art. Immunoprecipitations are performed using HSECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of HSECP expressed on the cell surface.

Alternatively, an assay for HSECP activity measures the amount of HSECP in secretory, membrane-bound organelles. Transfected cells as described above are harvested and lysed. The lysate is fractionated using methods known to those of skill in the art, for example, sucrose gradient ultracentrifugation. Such methods allow the isolation of subcellular components such as the Golgi apparatus, ER, small membrane-bound vesicles, and other secretory organelles.

Immunoprecipitations from fractionated and total cell lysates are performed using HSECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The concentration of HSECP in secretory organelles relative to HSECP in total cell lysate is proportional to the amount of HSECP in transit through the secretory pathway.

XI. Functional Assays

HSECP function is assessed by expressing the sequences encoding HSECP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1



plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HSECP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HSECP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HSECP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of HSECP Specific Antibodies

HSECP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HSECP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A



peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-HSECP activity by, for example, binding the peptide or HSECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring HSECP Using Specific Antibodies

Naturally occurring or recombinant HSECP is substantially purified by immunoaffinity chromatography using antibodies specific for HSECP. An immunoaffinity column is constructed by covalently coupling anti-HSECP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HSECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HSECP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HSECP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HSECP is collected.

XIV. Identification of Molecules Which Interact with HSECP

HSECP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HSECP, washed, and any wells with labeled HSECP complex are assayed. Data obtained using different concentrations of HSECP are used to calculate values for the number, affinity, and association of HSECP with the candidate molecules.

Alternatively, molecules interacting with HSECP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, *Nature* 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.



Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	23	078811	SYNORAB01	078811H1 (SYNORAB01), 078811P6 (SYNORAB01), 078811T6 (SYNORAB01),
2	24	371156	LUNGNOT02	077182F1 (SYNORAB01), 077182P1 (SYNORAB01), 289599H1 (TMLR3DT01), 371156H1 (LUNGNOT02), 523415H1 (MMLP2DT01), 3438708F6 (PENCNOT06), 3772179F6 (BRSTNOT25), 3898004H1 (CONTTUT01)
3	25	584050	PROSNOT02	584050H1 (PROSNOT02), 1726563T6 (PROSNOT14), 1739666R6 (HIPONON01), 1856214F6 (PROSNOT18), 2305379R6 (NGANNOT01), 2681374F6 (SINIUCT01), 4070575H1 (KIDNNOT26), 5274539H1 (OVARIN02), SZAI01719F1, SASA02714F1
4	26	863808	BRAITUT03	851821H1 (NGANNOT01), 863808H1 (BRAITUT03), 863808T1 (BRAITUT03), 2735728F6 (OVARNOT09)
5	27	978433	BRSTNOT02	054823R1 (FIBENOT01), 978433H1 (BRSTNOT02), 978433R1 (BRSTNOT02), 1867687T6 (SKINBIT01), 2503122H1 (CONUTUT01), 2522586H1 (GELANOT02), 3411659H1 (BRSTTUS08)
6	28	1655369	PROSTUT08	746013R1 (BRAITUT01), 944864F6 (FATRNOT02), 1539790R1 (SINTTUT01), 1617847F6 (BRAITUT12), 1655369F6 (PROSTUT08), 1655369H1 (PROSTUT08), 1673290F6 (BLANOT05), 2056840R6 (BEPINOT01), 3407992H1 (PROSTUS08), 4077365H1 (PANCNOT19), 4098012H1 (BRAITUT26)



Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
7	29	1703244	DUODNOT02	273878H1 (PANCIT03), 1632051H1 (COLNNOT19), 1703244H1 (DUODNOT02), 4176234H1 (SINTNOT21)
8	30	1730819	BRSTTUT08	862188R1 (BRAITUT03), 1399644F1 (BRAITUT08), 1443690F1 (THYRNOT03), 1596446F1 (BRAINOT14), 1730819F6 (BRSTTUT08), 1730819H1 (BRSTTUT08), 2304942H1 (NGANNOT01), 2868843H1 (THYFNOT10), 3395457H1 (LUNGNOT28)
9	31	1757161	PITUNOT03	864626R1 (BRAITUT03), 1231577H1 (BRAITUT01), 1395593F1 (THYRNOT03), 1466766T1 (PANCITUT02), 1466766T6 (PANCITUT02), 1597583F6 (BRAINOT14), 1757161H1 (PITUNOT03), 1757161R6 (PITUNOT03), 1757161T6 (PITUNOT03)
10	32	1976095	PANCITUT02	864976T1 (BRAITUT03), 1976095H1 (PANCITUT02), 4616101H1 (BRAYDIT01), SARB01143F1, SARB01861F1, SAJA02355F1
11	33	2169991	ENDCNOT03	2169991H1 (ENDCNOT03), 2313548R3 (NGANNOT01), 2727735T3 (OVARFUT05), 3095616H1 (CERVNOT03), 3186421H1 (THYMNOT04), 3493007H1 (ADPETUT07), SASA03601F1, SECA05288F1, SBDA00157F1
12	34	2616827	GBLANOT01	2616827H1 (GBLANOT01), 4637616F6 (MYEPTXT01), 5219252H1 (BRSTNOT35)
13	35	2991370	KIDUFET02	080447R1 (SYNOFAP01), 190292F1 (SYNOFAP01), 192279R1 (SYNOFAP01), 2724873H1 (OVARFUT05), 2991370H1 (KIDUFET02), SAUA03596F1, SAUA01525F1



Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
14	36	3031062	THYMNOT05	1354027T3 (LUNGNOT09), 2282215F6 (LUNGFEET04), 3031062H1 (THYMNOT05), SBAA04235F1, SBAA00620F1, SBAA01760F1, SBAA04365F2, SBAA01589F1
15	37	3101617	BRAINOT20	3101617H1 (BRAINOT20), 3335717F6 (BRAIFET01), SCAA04574V1, SCAA04351V1, SCAA03628V1, SCAA05459V1, SCAA01004V1
16	38	3216178	TESTNOT07	632084R6 (KIDNNOT05), 1993593T6 (CORPNOT02), 3216178F6 (TESTNOT07), 3216178H1 (TESTNOT07), 4914242H1 (LIVRFET05)
17	39	3406803	PROSTUS08	3406803F6 (PROSTUS08), 3406803H1 (PROSTUS08), 3406803T6 (PROSTUS08)
18	40	3468066	BRAIDIT01	659544H1 (BRAINOT03), 897142R1 (BPSTNOT05), 1321038F1 (BLADNOT04), 1351888F1 (LATRTUT02), 1485695F1 (CORPNOT02), 1507666F1 (LUNGNOT14), 2953291H1 (KIDNFET01), 3468066H1 (BRAIDIT01), 4426018F6 (BRAPDIT01)
19	41	3592862	293TF5T01	941610H1 (ADRENOT03), 1288036H1 (BRAINOT11), 1687969F6 (PROSTUT10), 3592862H1 (293TF5T01)
20	42	3669422	KIDNTUT16	3669422F6 (KIDNTUT16), 3669422H1 (KIDNTUT16), 3669422T6 (KIDNTUT16), 5445503H1 (LUNODNOT12)
21	43	3688740	HEAANOT01	462098R6 (LATPNOT01), 3688740H1 (HEAANOT01)
22	44	3742589	THYMNOT08	938795R1 (CERVNOT01), 1255960T1 (MENITUT03), 1530330R1 (PANCNOT04), 3742589H1 (THYMNOT08), SBOA04142D1



Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Potential Signal Peptides and Other Signature Sequences	Analytical Methods
1	182	S149 T151 S13		M1-S16; M1-P18	MOTIFS SPSCAN HMM
2	125	T87 S36 T78 T111	N75	M1-S37	MOTIFS SPSCAN HMM
3	320	S21 T63 S267 S300 T164	N40	M1-G20; M1-G23	MOTIFS SPSCAN HMM
4	234	T74 S198 T210 T227 S131 T195		M1-A30; M1-G25	MOTIFS SPSCAN HMM
5	278	S64 S132 S230 T252 S7 S179	N221	M1-A65	MOTIFS SPSCAN HMM
6	136	S98 S99		M1-P54	MOTIFS SPSCAN HMM
7	109	T57		M1-T20	MOTIFS SPSCAN HMM



Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Potential Signal Peptides and Other Signature Sequences	Analytical Methods
8	262	S73 S91 S136 S86	N182	M1-A25; M1-G27	MOTIFS SPSCAN HMM
9	384	S11 T140 S32 S185 S232 S306 S378	N50 N59 N62 N304	M1-G31 Transmembrane domains: G195-A220 L74-F91 A142-T160	MOTIFS SPSCAN HMM
10	244	T95 S110 T208 T44 S47 S53 S69 S152 T194		M1-A23	MOTIFS SPSCAN HMM
11	326	S138 T59 S239	N129 N237	M1-G40	MOTIFS SPSCAN HMM
12	105	S96 T24	N94	M1-A19; M1-A23	MOTIFS SPSCAN HMM



Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Potential Signal Peptides and Other Signature Sequences	Analytical Methods
13	626	S188 S593 T61 T118 T144 S252 T275 T410 S423 S501 T506 S524 T536 S550 S28 T146 S336 T467 S583 T588		M1-G27 RGD tripeptide: R271-D273	MOTIFS SPSCAN HMM
14	296	S205 T295 S109 T165 T214 S244 S73 S225 Y236	N40 N53 N204 N281	M1-G23	MOTIFS SPSCAN HMM
15	249	S7 T123 S233 S237 T151	N100	M1-G23; M1-A22 Transmembrane domains: T175-A193 P97-F120	MOTIFS SPSCAN HMM
16	124	S62 S109		M1-G37 Somatomedin B signature: L36-S105	MOTIFS SPSCAN HMM PROFILES SCAN
17	101	T59 S67		M1-S40; M1-S28	MOTIFS SPSCAN HMM



Table 2 (cont.)

Polyptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Potential Signal Peptides and Other Signature Sequences	Analytical Methods
18	540	S42 S38 T80 T172 S179 S326 S519 S531 T131 T250 T278 Y99	N78 N88 N170 N347 N448 N457	M1-C29; M1-L31 Transmembrane domains: V307-N327 L411-I428 A140-L163 D366-I385 Y99-Y122 F496-C513 I56-Y73	MOTIFS SPSCAN HMM PRINTS
19	108	S34 S64 S29 S47 Y96		M1-R22; M1-S23	MOTIFS SPSCAN HMM
20	114	S108 S5 S73 T85		M1-T32	MOTIFS SPSCAN HMM
21	114	T8	N6	M1-A34	MOTIFS SPSCAN HMM
22	287	S25 T75 T134 T139 S225 T255 S254		M1-P20; M1-A22 RGD tripeptide: R172-D174	MOTIFS SPSCAN HMM



Table 3

Nucleotide SEQ ID NO:	Selected Fragment(s)	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
23	110-154 542-586	Musculoskeletal (0.750) Reproductive (0.250)	Inflammation (0.750) Cancer (0.250)	PBLUESCRIPT
24	230-271	Reproductive (0.333) Musculoskeletal (0.190) Cardiovascular (0.143)	Cancer (0.476) Inflammation (0.333)	PBLUESCRIPT
25	109-153 649-693	Reproductive (0.314) Nervous (0.235) Gastrointestinal (0.157)	Cancer (0.608) Inflammation (0.196) Cell Proliferation (0.118)	PSPORT1
26	116-160	Reproductive (0.333) Nervous (0.178) Cardiovascular (0.156)	Cancer (0.600) Cell Proliferation (0.178) Inflammation (0.133)	PSPORT1
27	228-272	Reproductive (0.333) Cardiovascular (0.244) Gastrointestinal (0.111)	Cancer (0.667) Cell Proliferation (0.133) Inflammation (0.089)	PSPORT1
28	1945-1989	Nervous (0.301) Reproductive (0.219) Gastrointestinal (0.137)	Cancer (0.452) Inflammation (0.205) Trauma (0.164)	pINCY
29	271-315	Gastrointestinal (0.800) Nervous (0.200)	Trauma (0.600) Cancer (0.200) Inflammation (0.200)	pINCY
30	218-262	Reproductive (0.249) Nervous (0.195) Gastrointestinal (0.136)	Cancer (0.521) Inflammation (0.207) Cell Proliferation (0.172)	pINCY



Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragment(s)	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
31	1190-1234	Nervous (0.324) Reproductive (0.250) Hematopoietic/Immune (0.118)	Cancer (0.426) Inflammation (0.279) Cell Proliferation (0.147)	PSPORT1
32	487-531	Gastrointestinal (0.333) Nervous (0.259) Reproductive (0.111)	Cancer (0.593) Inflammation (0.148) Cell Proliferation (0.111)	pINCY
33	1513-1557	Reproductive (0.255) Nervous (0.216) Hematopoietic/Immune (0.176)	Cancer (0.373) Inflammation (0.255) Cell Proliferation (0.196)	pINCY
34	270-317	Hematopoietic/Immune (0.455) Musculoskeletal (0.182) Nervous (0.182)	Inflammation (0.545) Cancer (0.273) Cell Proliferation (0.091)	pINCY
35	1299-1343 1956-2000	Musculoskeletal (0.519) Nervous (0.148) Cardiovascular (0.111)	Inflammation (0.481) Cancer (0.296) Trauma (0.074)	pINCY
36	651-695	Nervous (0.250) Developmental (0.167) Gastrointestinal (0.167)	Cell Proliferation (0.667) Inflammation (0.250) Cancer (0.167)	pINCY
37	218-262	Nervous (0.818) Gastrointestinal (0.091) Reproductive (0.091)	Cancer (0.545) Cell Proliferation (0.091) Inflammation (0.091)	pINCY
38	290-334	Hematopoietic/Immune (0.250) Urologic (0.250) Developmental (0.125)	Cell Proliferation (0.375) Inflammation (0.250) Cancer (0.125)	pINCY



Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragment(s)	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
39	434-478	Reproductive (1.000)	Cancer (1.000)	pINCY
40	326-370	Nervous (0.242) Reproductive (0.220) Gastrointestinal (0.121)	Cancer (0.462) Inflammation (0.280) Cell Proliferation (0.121)	pINCY
41	165-209	Nervous (0.333) Gastrointestinal (0.200) Cardiovascular (0.133)	Cancer (0.467) Cell Proliferation (0.200) Inflammation (0.133)	pINCY
42	273-317	Hematopoietic/Immune (0.312) Nervous (0.188) Reproductive (0.167)	Cancer (0.354) Inflammation (0.312) Cell Proliferation (0.146)	pINCY
43	273-317	Cardiovascular (1.000)	Inflammation (0.500)	pINCY
44	435-479	Reproductive (0.297) Nervous (0.217) Gastrointestinal (0.109)	Cancer (0.464) Cell Proliferation (0.196) Inflammation (0.167)	pINCY



Table 4

Nucleotide SEQ ID NO:	Library	Library Description
23	SYNORAB01	This library was constructed using RNA isolated from the synovial membrane tissue of a 68-year-old Caucasian female with rheumatoid arthritis.
24	LUNGNOT02	This library was constructed using RNA isolated from the lung tissue of a 47-year-old Caucasian male who died of a subarachnoid hemorrhage.
25	PROSNOT02	This library was constructed using RNA isolated from diseased prostate tissue removed from a 50-year-old Caucasian male during a retropubic prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma Gleason grade 3+3. Patient history included dysuria, carcinoma in situ of prostate, coronary atherosclerosis, and hyperlipidemia.
26	BRAITUT03	This library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 17-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
27	BRSTNOT02	This library was constructed using RNA isolated from diseased breast tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated proliferative fibrocystic changes characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Pathology for the associated tumor tissue indicated an invasive grade 4 mammary adenocarcinoma. Patient history included atrial tachycardia and a benign neoplasm. Family history included cardiovascular and cerebrovascular disease.



Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
28	PROSTUT08	This library was constructed using RNA isolated from prostate tumor tissue removed from a 60-year-old Caucasian male during radical prostatectomy and regional lymph node excision. Pathology indicated an adenocarcinoma (Gleason grade 3+4). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Patient history included a kidney cyst, and hematuria. Family history included tuberculosis, cerebrovascular disease, and arteriosclerotic coronary artery disease.
29	DUODNOT02	This library was constructed using RNA isolated from duodenal tissue of an 8-year-old Caucasian female, who died from head trauma. Serology was positive for cytomegalovirus (CMV).
30	BRSTTUT08	This library was constructed using RNA isolated from breast tumor tissue removed from a 45-year-old Caucasian female during unilateral extended simple mastectomy. Pathology indicated invasive nuclear grade 2-3 adenocarcinoma, ductal type, with 3 of 23 lymph nodes positive for metastatic disease. Greater than 50% of the tumor volume was in situ, both comedo and non-comedo types. Immunostains were positive for estrogen/progesterone receptors, and uninvolved tissue showed proliferative changes. The patient concurrently underwent a total abdominal hysterectomy. Patient history included valvuloplasty of mitral valve without replacement, rheumatic mitral insufficiency, and rheumatic heart disease. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
31	PITUNOT03	This library was constructed using RNA isolated from pituitary tissue of a 46-year-old Caucasian male who died from colon cancer.
32	PANCTUT02	This library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreaticoduodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.



Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
33	ENDCNOT03	This library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a neonatal Caucasian male.
34	GBLANOT01	This library was constructed using RNA isolated from diseased gallbladder tissue removed from a 53-year-old Caucasian female during a cholecystectomy. Pathology indicated mild chronic cholecystitis and cholelithiasis with approximately 150 mixed gallstones. Family history included benign hypertension.
35	KIDNFET02	This library was constructed using RNA isolated from kidney tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
36	TLYMNOT05	This library was constructed using RNA isolated from nonactivated Th2 cells. The cells were differentiated from umbilical cord CD4 T cells with IL-4 in the presence of anti-IL-12 antibodies and B7-transfected COS cells.



Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
37	BRAINOT20	This library was constructed using RNA isolated from diseased brain tissue removed from the left temporal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology for the left temporal lobe, including the mesial temporal structures, indicated focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. The left frontal lobe showed a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This frontal lobe tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. GFAP was positive for astrocytes. Family history included brain cancer.
38	TESTNOT07	This library was constructed using RNA isolated from testicular tissue removed from a 31-year-old Caucasian male during a unilateral orchiectomy (excision of testis). Pathology indicated a mass containing a large subcapsular hematoma with laceration of the tunica albuginea. The surrounding testicular parenchyma was extensively necrotic. The patient presented with a trunk injury.
39	PROSTUS08	This subtracted library was constructed using 2.36 million clones from a prostate tumor library and was subjected to one round of subtractive hybridization with 448,000 clones from a prostate library. The starting library for subtraction was constructed using RNA isolated from a prostate tumor removed from a 59-year-old Caucasian male during a radical prostatectomy with regional lymph node excision. Pathology indicated adenocarcinoma (Gleason grade 3+3) and adenofibromatous hyperplasia. The patient presented with elevated prostate-specific antigen (PSA). Patient history included colon diverticuli, asbestosis, and thrombophlebitis. Family history included multiple myeloma, hyperlipidemia, and rheumatoid arthritis. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR (1991) 19:1954 and Ronaldo, et al. Genome Research (1996) 6:791.
40	BRAIDIT01	This library was constructed using RNA isolated from diseased brain tissue. Patient history included multiple sclerosis, type II lesion.



Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
41	293TF5T01	This library was constructed using RNA isolated from a transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue transfected with bgal. The cells were transformed with adenovirus 5 DNA.
42	KIDNTUT16	This library was constructed using RNA isolated from left pole kidney tumor tissue removed from a 53-year-old Caucasian female during a nephroureterectomy. Pathology indicated grade 2 renal cell carcinoma. Patient history included hyperlipidemia, cardiac dysrhythmia, menorrhagia, cerebrovascular disease, atherosclerotic coronary artery disease, and tobacco abuse. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
43	HEANOT01	This library was constructed using RNA isolated from right coronary and right circumflex coronary artery tissue removed from the explanted heart of a 46-year-old Caucasian male during a heart transplantation. Patient history included myocardial infarction from total occlusion of the left anterior descending coronary artery, atherosclerotic coronary artery disease, hyperlipidemia, myocardial ischemia, dilated cardiomyopathy, left ventricular dysfunction, and tobacco abuse. Previous surgeries included cardiac catheterization. Family history included atherosclerotic coronary artery disease.
44	THYMNOT08	This library was constructed using RNA isolated from thymus tissue removed from a 4-month-old Caucasian male during a total thymectomy and open heart repair of atrioventricular canal defect using hypothermia. Patient presented with a congenital heart anomaly, congestive heart failure, and Down syndrome. Patient history included abnormal thyroid function study, premature birth, and right and left heart angiocardiology.



Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/FARACEL EDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families



Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score \geq GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=14-21
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <i>supra</i> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	



What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

- 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-22,
 b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-22,
 c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22, and
10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-22.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-22.

15

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide of claim 3 selected from the group consisting of SEQ ID NO:23-44.

20

5. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

6. A cell transformed with a recombinant polynucleotide of claim 5.

25

7. A transgenic organism comprising a recombinant polynucleotide of claim 5.

8. A method for producing a polypeptide of claim 1, the method comprising:

- 30 a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
 b) recovering the polypeptide so expressed.

35

9. An isolated antibody which specifically binds to a polypeptide of claim 1.



10. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44.
- b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:23-44.
- c) a polynucleotide sequence complementary to a).
- d) a polynucleotide sequence complementary to b), and
- e) an RNA equivalent of a)-d).

11. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 10.

12. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 10, the method comprising:

- a) hybridizing the sample with a probe comprising at least 16 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

13. A method of claim 12, wherein the probe comprises at least 30 contiguous nucleotides.

14. A method of claim 12, wherein the probe comprises at least 60 contiguous nucleotides.

15. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

16. A method for treating a disease or condition associated with decreased expression of functional HSECP, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 15.

17. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and



b) detecting agonist activity in the sample.

18. A pharmaceutical composition comprising an agonist compound identified by a method of claim 17 and a pharmaceutically acceptable excipient.

5

19. A method for treating a disease or condition associated with decreased expression of functional HSECP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 18.

10 20. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

15 21. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

20 22. A method for treating a disease or condition associated with overexpression of functional HSECP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 21.

23. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 4, the method comprising:

- 25
- a) exposing a sample comprising the target polynucleotide to a compound, and
 - b) detecting altered expression of the target polynucleotide.



SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.

TANG, Y. Tom
 LAL, Preeti
 BAUGHN, Mariah R.
 YUE, Henry
 AU-YOUNG, Janice
 LU, Dyung Aina M.
 AZIMZAI, Valda

<120> HUMAN SECRETORY PROTEINS

<130> PF-0675 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/123,117

<151> 1999-03-05

<160> 44

<170> PERL Program

<210> 1

<211> 182

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 078811CD1

<400> 1

Met	Arg	Ser	Thr	Ile	Leu	Leu	Phe	Cys	Leu	Leu	Gly	Ser	Thr	Arg			
1				5					10					15			
Ser	Leu	Pro	Val	Phe	Pro	Ser	Leu	Ser	Leu	Ile	Pro	Leu	Thr	Gln			
				20					25					30			
Met	Leu	Thr	Leu	Gly	Pro	Asp	Leu	His	Leu	Leu	Asn	Pro	Ala	Ala			
				35					40					45			
Gly	Met	Thr	Pro	Gly	Thr	Gln	Thr	His	Pro	Leu	Thr	Leu	Gly	Gly			
				50					55					60			
Leu	Asn	Val	Gln	Gln	Gln	Leu	His	Pro	His	Val	Leu	Pro	Ile	Phe			
				65					70					75			
Val	Thr	Gln	Leu	Gly	Ala	Pro	Gly	His	Tyr	Pro	Lys	Leu	Arg	Gly			
				80					85					90			
Ile	Ala	Thr	Asn	Leu	His	Glu	Pro	His	His	Pro	Phe	Leu	Val	Pro			
				95					100					105			
Arg	Glu	Ala	Ser	Leu	Pro	Thr	Ser	Gln	Ala	Gly	Ala	Asn	Pro	Asp			
				110					115					120			
Val	Gln	Asp	Gly	Ser	Leu	Pro	Ala	Gly	Gly	Ala	Gly	Val	Asn	Pro			
				125					130					135			
Ala	Thr	Gln	Gly	Thr	Pro	Ala	Gly	Arg	Leu	Pro	Thr	Pro	Ser	Gly			
				140					145					150			
Thr	Asp	Asp	Asp	Phe	Ala	Val	Thr	Thr	Pro	Ala	Gly	Ile	Gln	Arg			
				155					160					165			
Ser	Thr	His	Ala	Ile	Glu	Glu	Ala	Thr	Thr	Glu	Ser	Ala	Asn	Gly			
				170					175					180			
Ile	Gln																

<210> 2

<211> 125



<212> PRT
 <213> Homo sapiens
 <220> -
 <221> misc_feature
 <223> Incyte ID No: 371156CD1

<400> 2
 Met Val Cys Glu Asp Ala Pro Ser Phe Gln Met Ala Trp Glu Ser
 1 5 10 15
 Gln Met Ala Trp Glu Arg Gly Pro Ala Leu Leu Cys Cys Val Leu
 20 25 30
 Ser Ala Ser Gln Leu Ser Ser Gln Asp Gln Asp Pro Leu Gly His
 35 40 45
 Ile Lys Ser Leu Leu Tyr Pro Phe Gly Phe Pro Val Glu Leu Pro
 50 55 60
 Arg Pro Gly Pro Thr Gly Ala Tyr Lys Lys Val Lys Asn Gln Asn
 65 70 75
 Gln Thr Thr Ser Ser Glu Leu Leu Arg Lys Gln Thr Ser His Phe
 80 85 90
 Asn Gln Arg Gly His Arg Ala Arg Ser Lys Leu Leu Ala Ser Arg
 95 100 105
 Gln Ile Pro Asp Arg Thr Phe Lys Cys Gly Lys Trp Leu Pro Gln
 110 115 120
 Val Pro Ser Pro Val
 125

<210> 3
 <211> 320
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 584050CD1

<400> 3
 Met Ala Gly Leu Ala Ala Arg Leu Val Leu Leu Ala Gly Ala Ala
 1 5 10 15
 Ala Leu Ala Ser Gly Ser Gln Gly Asp Arg Glu Pro Val Tyr Arg
 20 25 30
 Asp Cys Val Leu Gln Cys Glu Glu Gln Asn Cys Ser Gly Gly Ala
 35 40 45
 Leu Asn His Phe Arg Ser Arg Gln Pro Ile Tyr Met Ser Leu Ala
 50 55 60
 Gly Trp Thr Cys Arg Asp Asp Cys Lys Tyr Glu Cys Met Trp Val
 65 70 75
 Thr Val Gly Leu Tyr Leu Gln Glu Gly His Lys Val Pro Gln Phe
 80 85 90
 His Gly Lys Trp Pro Phe Ser Arg Phe Leu Phe Phe Gln Glu Pro
 95 100 105
 Ala Ser Ala Val Ala Ser Phe Leu Asn Gly Leu Ala Ser Leu Val
 110 115 120
 Met Leu Cys Arg Tyr Arg Thr Phe Val Pro Ala Ser Ser Pro Met
 125 130 135
 Tyr His Thr Cys Val Ala Phe Ala Trp Val Ser Leu Asn Ala Trp
 140 145 150
 Phe Trp Ser Thr Val Phe His Thr Arg Asp Thr Asp Leu Thr Glu
 155 160 165
 Lys Met Asp Tyr Phe Cys Ala Ser Thr Val Ile Leu His Ser Ile
 170 175 180
 Tyr Leu Cys Cys Val Arg Thr Val Gly Leu Gln His Pro Ala Val
 185 190 195
 Val Ser Ala Phe Arg Ala Leu Leu Leu Met Leu Thr Val His



Val Ser Tyr Leu	200	Ser Leu Ile Arg Phe	205	Asp Tyr Gly Tyr Asn	210
	215		220		225
Val Ala Asn Val	230	Ala Ile Gly Leu Val	235	Asn Val Val Trp Trp	240
	245		250		255
Ala Trp Cys Leu	260	Trp Asn Gln Arg Arg	265	Leu Pro His Val Arg	270
	275		280		285
Cys Val Val Val	290	Val Leu Leu Leu Gln	295	Gly Leu Ser Leu Leu	300
	305		310		315
Leu Leu Asp Phe	320	Pro Pro Leu Phe Trp		Val Leu Asp Ala His	
Ile Trp His Ile		Ser Thr Ile Pro Val		His Val Leu Phe Phe	
Phe Leu Glu Asp		Asp Ser Leu Tyr Leu		Leu Lys Glu Ser Glu	
Lys Phe Lys Leu					

<210> 4

<211> 234

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 863808CD1

<400> 4

Met Gly Pro Gly Gly	Arg Val Ala Arg	Leu Leu Ala Pro Leu	Met
1	5	10	15
Trp Arg Arg Ala Val	Ser Ser Val Ala	Gly Ser Ala Val Gly	Ala
	20	25	30
Glu Pro Gly Leu Arg	Leu Leu Ala Val	Gln Arg Leu Pro Val	Gly
	35	40	45
Ala Ala Phe Cys Arg	Ala Cys Gln Thr	Pro Asn Phe Val Arg	Gly
	50	55	60
Leu His Ser Glu Pro	Gly Leu Glu Glu	Arg Ala Glu Gly Thr	Val
	65	70	75
Asn Glu Gly Arg Pro	Glu Ser Asp Ala	Ala Asp His Thr Gly	Pro
	80	85	90
Lys Phe Asp Ile Asp	Met Met Val Ser	Leu Leu Arg Gln Glu	Asn
	95	100	105
Ala Arg Asp Ile Cys	Val Ile Gln Val	Pro Pro Glu Met Arg	Tyr
	110	115	120
Thr Asp Tyr Phe Val	Ile Val Ser Gly	Thr Ser Thr Arg His	Leu
	125	130	135
His Ala Met Ala Phe	Tyr Val Val Lys	Met Tyr Lys His Leu	Lys
	140	145	150
Cys Lys Arg Asp Pro	His Val Lys Ile	Glu Gly Lys Asp Thr	Asp
	155	160	165
Asp Trp Leu Cys Val	Asp Phe Gly Ser	Met Val Ile His Leu	Met
	170	175	180
Leu Pro Glu Thr Arg	Glu Ile Tyr Glu	Leu Glu Lys Leu Trp	Thr
	185	190	195
Leu Arg Ser Tyr Asp	Asp Gln Leu Ala	Gln Ile Ala Pro Glu	Thr
	200	205	210
Val Pro Glu Asp Phe	Ile Leu Gly Ile	Glu Asp Asp Thr Ser	Ser
	215	220	225
Val Thr Pro Val Glu	Leu Lys Cys Glu		
	230		



<210> 5
 <211> 278
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 978433CD1

<400> 5
 Met Gln Pro Ala Ala Ala Ser Glu Arg Gly Gly Ala Asp Ala Asp
 1 5 10 15
 His Val Pro Leu Leu Gly Leu Leu Arg Leu Gln Leu Arg Ala Ala
 20 25 30
 Arg Gln Pro Gly Ala Met Arg Pro Gln Gly Pro Ala Ala Ser Pro
 35 40 45
 Gln Arg Leu Arg Gly Leu Leu Leu Leu Leu Leu Gln Leu Pro
 50 55 60
 Ala Pro Ser Ser Ala Ser Glu Ile Pro Lys Gly Lys Gln Lys Ala
 65 70 75
 Gln Leu Arg Gln Arg Glu Val Val Asp Leu Tyr Asn Gly Met Cys
 80 85 90
 Leu Gln Gly Pro Ala Gly Val Pro Gly Arg Asp Gly Ser Pro Gly
 95 100 105
 Ala Asn Gly Ile Pro Gly Thr Pro Gly Ile Pro Gly Arg Asp Gly
 110 115 120
 Phe Lys Gly Glu Lys Gly Glu Cys Leu Arg Glu Ser Phe Glu Glu
 125 130 135
 Ser Trp Thr Pro Asn Tyr Lys Gln Cys Ser Trp Ser Ser Leu Asn
 140 145 150
 Tyr Gly Ile Asp Leu Gly Lys Ile Ala Glu Cys Thr Phe Thr Lys
 155 160 165
 Met Arg Ser Asn Ser Ala Leu Arg Val Leu Phe Ser Gly Ser Leu
 170 175 180
 Arg Leu Lys Cys Arg Asn Ala Cys Cys Gln Arg Trp Tyr Phe Thr
 185 190 195
 Phe Asn Gly Ala Glu Cys Ser Gly Pro Leu Pro Ile Glu Ala Ile
 200 205 210
 Ile Tyr Leu Asp Gln Gly Ser Pro Glu Met Asn Ser Thr Ile Asn
 215 220 225
 Ile His Arg Thr Ser Ser Val Glu Gly Leu Cys Glu Gly Ile Gly
 230 235 240
 Ala Gly Leu Val Asp Val Ala Ile Trp Val Gly Thr Cys Ser Asp
 245 250 255
 Tyr Pro Lys Gly Asp Ala Ser Thr Gly Trp Asn Ser Val Ser Arg
 260 265 270
 Ile Ile Ile Glu Glu Leu Pro Lys
 275

<210> 6
 <211> 136
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1655369CD1

<400> 6
 Met Pro Pro Gly Gly Leu Gly Ala Cys Ala Val Thr Pro Ala Pro
 1 5 10 15
 Gly Glu Glu Arg Thr Gln Pro Gly Glu Leu Gly Gln Gly Leu His
 20 25 30
 Met Ala Gln Gly Cln Cln Met Leu Ala Gly Gln Leu Leu Pro Met



Leu Thr Leu Leu	35	Pro Ser Phe Pro	40	Pro His Pro Thr	45
	50		55		60
Gly Pro Arg Arg	65	Ala Ser Leu Thr	70	Leu Gly Pro Ala	75
	80		85		90
Trp Met Ala Trp	95	Arg Pro Trp Ala	100	Leu Gly Pro Gly	105
	110		115		120
Pro Leu Gly Gln	125	Trp Lys Ser Ser	130	Val Glu Glu His	135
Ala Ala Trp Leu		Pro Leu Ala Leu		Glu Trp Ser Leu	
Ala Ser Ala Leu		Ala Leu Gly Thr		Ser His Pro Leu	
Gln					

<110> 7
 <111> 109
 <112> PRT
 <113> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1703244CD1

Met Leu Met Tyr	Met	Phe Tyr Val	Leu	Pro	Phe Cys Gly	Leu	Ala
	5			10			15
Ala Tyr Ala Leu	Thr	Phe Pro Gly	Cys	Ser	Trp Leu Pro	Asp	Trp
	20			25			30
Ala Leu Val Phe	Ala	Gly Gly Ile	Gly	Gln	Ala Gln Phe	Ser	His
	35			40			45
Met Gly Ala Ser	Met	His Leu Arg	Thr	Pro	Phe Thr Tyr	Arg	Val
	50			55			60
Pro Glu Asp Thr	Trp	Gly Cys Phe	Phe	Val	Cys Asn Leu	Leu	Tyr
	65			70			75
Ala Leu Gly Pro	His	Leu Leu Ala	Tyr	Arg	Cys Leu Gln	Trp	Pro
	80			85			90
Ala Phe Phe His	Gln	Pro Pro Pro	Ser	Asp	Pro Leu Ala	Leu	His
	95			100			105
Lys Lys Gln His							

<210> 8
 <211> 262
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1730819CD1

Met Ala Ala Ala	Ser	Ala Gly Ala	Thr	Arg	Leu Leu Leu	Leu	Leu
	5			10			15
Leu Met Ala Val	Ala	Ala Pro Ser	Arg	Ala	Arg Gly Ser	Gly	Cys
	20			25			30
Arg Ala Gly Thr	Gly	Ala Arg Gly	Ala	Gly	Ala Glu Gly	Arg	Glu
	35			40			45
Gly Glu Ala Cys	Gly	Thr Val Gly	Leu	Leu	Leu Glu His	Ser	Phe
	50			55			60
Glu Ile Asp Asp	Ser	Ala Asn Phe	Arg	Lys	Arg Gly Ser	Leu	Leu
	65			70			75



Trp	Asn	Gln	Gln	Asp	Gly	Thr	Leu	Ser	Leu	Ser	Gln	Arg	Gln	Leu	
				80					85					90	
Ser	Glu	Glu	Glu	Arg	Gly	Arg	Leu	Arg	Asp	Val	Ala	Ala	Leu	Asn	
				95					100					105	
Gly	Leu	Tyr	Arg	Val	Arg	Ile	Pro	Arg	Arg	Pro	Gly	Ala	Leu	Asp	
				110					115					120	
Gly	Leu	Glu	Ala	Gly	Gly	Tyr	Val	Ser	Ser	Phe	Val	Pro	Ala	Cys	
				125					130					135	
Ser	Leu	Val	Glu	Ser	His	Leu	Ser	Asp	Gln	Leu	Thr	Leu	His	Val	
				140					145					150	
Asp	Val	Ala	Gly	Asn	Val	Val	Gly	Val	Ser	Val	Val	Thr	His	Pro	
				155					160					165	
Gly	Gly	Cys	Arg	Gly	His	Glu	Val	Glu	Asp	Val	Asp	Leu	Glu	Leu	
				170					175					180	
Phe	Asn	Thr	Ser	Val	Gln	Leu	Gln	Pro	Pro	Thr	Thr	Ala	Pro	Gly	
				185					190					195	
Pro	Glu	Thr	Ala	Ala	Phe	Ile	Glu	Arg	Leu	Glu	Met	Glu	Gln	Ala	
				200					205					210	
Gln	Lys	Ala	Lys	Asn	Pro	Gln	Glu	Gln	Lys	Ser	Phe	Phe	Ala	Lys	
				215					220					225	
Tyr	Trp	Met	Tyr	Ile	Ile	Pro	Val	Val	Leu	Phe	Leu	Met	Met	Ser	
				230					235					240	
Gly	Ala	Pro	Asp	Thr	Gly	Gly	Gln	Gly	Gly	Gly	Gly	Gly	Cys	Gly	
				245					250					255	
Gly	Gly	Gly	Gly	Ser	Gly	Arg									
				250											

<210> 9

<211> 384

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1757161CD1

<400> 9

Met	Ala	Glu	Gln	Thr	Tyr	Ser	Trp	Ala	Tyr	Ser	Leu	Val	Asp	Ser	
1				5					10					15	
Ser	Gln	Val	Ser	Thr	Phe	Leu	Ile	Ser	Ile	Leu	Leu	Ile	Val	Tyr	
				20					25					30	
Gly	Ser	Phe	Arg	Ser	Leu	Asn	Met	Asp	Phe	Glu	Asn	Gln	Asp	Lys	
				35					40					45	
Glu	Lys	Asp	Ser	Asn	Ser	Ser	Ser	Gly	Ser	Phe	Asn	Gly	Asn	Ser	
				50					55					60	
Thr	Asn	Asn	Ser	Ile	Gln	Thr	Ile	Asp	Ser	Thr	Gln	Ala	Leu	Phe	
				65					70					75	
Leu	Pro	Ile	Gly	Ala	Ser	Val	Ser	Leu	Leu	Val	Met	Phe	Phe	Phe	
				80					85					90	
Phe	Asp	Ser	Val	Gln	Val	Val	Phe	Thr	Ile	Cys	Thr	Ala	Val	Leu	
				95					100					105	
Ala	Thr	Ile	Ala	Phe	Ala	Phe	Leu	Leu	Leu	Pro	Met	Cys	Gln	Tyr	
				110					115					120	
Leu	Thr	Arg	Pro	Cys	Ser	Pro	Gln	Asn	Lys	Ile	Ser	Phe	Gly	Cys	
				125					130					135	
Cys	Gly	Arg	Phe	Thr	Ala	Ala	Glu	Leu	Leu	Ser	Phe	Ser	Leu	Ser	
				140					145					150	
Val	Met	Leu	Val	Leu	Ile	Trp	Val	Leu	Thr	Gly	His	Trp	Leu	Leu	
				155					160					165	
Met	Asp	Ala	Leu	Ala	Met	Gly	Leu	Cys	Val	Ala	Met	Ile	Ala	Phe	
				170					175					180	
Val	Arg	Leu	Pro	Ser	Leu	Lys	Val	Ser	Cys	Leu	Leu	Leu	Ser	Gly	
				185					190					195	
Leu	Leu	Ile	Tyr	Asp	Val	Phe	Trp	Val	Phe	Phe	Ser	Ala	Tyr	Ile	



Phe Asn Ser Asn Val	200	Met Val Lys Val	205	Thr Gln Pro Ala Asp	210
215	220	225	230	235	240
Asn Pro Leu Asp Val	245	Leu Ser Arg Lys	250	Leu His Leu Gly Pro Asn	255
260	265	270	275	280	285
Val Gly Arg Asp Val	290	Pro Arg Leu Ser	295	Leu Pro Gly Lys Leu Val	300
305	310	315	320	325	330
Phe Pro Ser Ser Thr	335	Gly Ser His Phe	340	Ser Met Leu Gly Ile Gly	345
350	355	360	365	370	375
Asp Ile Val Met Pro	380	Gly Leu Leu Leu	Cys Phe Val Leu Arg Tyr		
385	390	395	400	405	410
Asp Asn Tyr Lys Lys	415	Gln Ala Ser Gly	420	Asp Ser Cys Gly Ala Pro	425
430	435	440	445	450	455
Gly Pro Ala Asn Ile	460	Ser Gly Arg Met	465	Gln Lys Val Ser Tyr Phe	470
475	480	485	490	495	500
His Cys Thr Leu Ile	505	Gly Tyr Phe Val	510	Gly Leu Leu Thr Ala Thr	515
520	525	530	535	540	545
Val Ala Ser Arg Ile	550	His Arg Ala Ala	555	Gln Pro Ala Leu Leu Tyr	560
565	570	575	580	585	590
Leu Val Pro Phe Thr	595	Leu Leu Pro Leu	600	Leu Thr Met Ala Tyr Leu	605
610	615	620	625	630	635
Lys Gly Asp Leu Arg	640	Arg Met Trp Ser	645	Glu Pro Phe His Ser Lys	650
655	660	665	670	675	680
Ser Ser Ser Ser Arg	685	Phe Leu Glu Val			
690	695	700	705	710	715

<210> 10

<211> 244

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1976095CD1

<400> 10

Met Asp Ile Leu Val	1	Pro Leu Leu Gln	10	Leu Leu Val Leu Leu	15
1	5	10	15	20	25
Thr Leu Pro Leu His	30	Leu Met Ala Leu	35	Gly Cys Trp Gln Pro	40
40	45	50	55	60	65
Leu Cys Lys Ser Tyr	70	Phe Pro Tyr Leu	75	Met Ala Val Leu Thr	80
80	85	90	95	100	105
Lys Ser Asn Arg Lys	110	Met Glu Ser Lys	115	Arg Glu Leu Phe Ser	120
120	125	130	135	140	145
Gln Ile Lys Gly Leu	150	Thr Gly Ala Ser	155	Gly Lys Val Ala Leu	160
160	165	170	175	180	185
Glu Leu Gly Cys Gly	190	Thr Gly Ala Asn	195	Phe Gln Phe Tyr Pro	200
200	205	210	215	220	225
Gly Cys Arg Val Thr	230	Cys Leu Asp Pro	235	Asn Pro His Phe Glu	240
240	245	250	255	260	265
Phe Leu Thr Lys Ser	270	Met Ala Glu Asn	275	Arg His Leu Gln Tyr	280
280	285	290	295	300	305
Arg Phe Val Val Ala	310	Pro Gly Glu Asp	315	Met Arg Gln Leu Ala	320
320	325	330	335	340	345
Gly Ser Met Asp Val	350	Val Val Cys Thr	355	Leu Val Leu Cys Ser	360
360	365	370	375	380	385
Gln Ser Pro Arg Lys	390	Val Leu Gln Glu	395	Val Arg Arg Val Leu	400
400	405	410	415	420	425
Pro Gly Gly Val Leu	430	Phe Phe Trp Glu	435	Val Ala Glu Pro Tyr	440
440	445	450	455	460	465
Gly Ser Trp Ala Phe	470	Met Trp Gln Gln	475	Val Phe Glu Pro Thr	480
480	485	490	495	500	505
Lys His Ile Gly Asp	510	Gly Cys Cys Leu	515	Thr Arg Glu Thr Trp	520
520	525	530	535	540	545



Asp	Leu	Glu	Asn	Ala	Gln	Phe	Ser	Glu	Ile	Gln	Met	Glu	Arg	Gln
				215					220					225
Pro	Pro	Pro	Leu	Lys	Trp	Leu	Pro	Val	Gly	Pro	His	Ile	Met	Gly
				230					235					240
Lys	Ala	Val	Lys											

<210> 11
 <211> 326
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2169991CD1

<400> 11

Met	Arg	Thr	Glu	Ala	Gln	Val	Pro	Ala	Leu	Gln	Pro	Pro	Glu	Pro
1				5					10					15
Gly	Leu	Glu	Gly	Ala	Met	Gly	His	Arg	Thr	Leu	Val	Leu	Pro	Trp
				20					25					30
Val	Leu	Leu	Thr	Leu	Cys	Val	Thr	Ala	Gly	Thr	Pro	Glu	Val	Trp
				35					40					45
Val	Gln	Val	Arg	Met	Glu	Ala	Thr	Glu	Leu	Ser	Ser	Phe	Thr	Ile
				50					55					60
Arg	Cys	Gly	Phe	Leu	Gly	Ser	Gly	Ser	Ile	Ser	Leu	Val	Thr	Val
				65					70					75
Ser	Trp	Gly	Gly	Pro	Asn	Gly	Ala	Gly	Gly	Thr	Thr	Leu	Ala	Val
				80					85					90
Leu	His	Pro	Glu	Arg	Gly	Ile	Arg	Gln	Trp	Ala	Pro	Ala	Arg	Gln
				95					100					105
Ala	Arg	Trp	Glu	Thr	Gln	Ser	Ser	Ile	Ser	Leu	Ile	Leu	Glu	Gly
				110					115					120
Ser	Gly	Ala	Ser	Ser	Pro	Cys	Ala	Asn	Thr	Thr	Phe	Cys	Cys	Lys
				125					130					135
Phe	Ala	Ser	Phe	Pro	Glu	Gly	Ser	Trp	Glu	Ala	Cys	Gly	Ser	Leu
				140					145					150
Pro	Pro	Ser	Ser	Asp	Pro	Gly	Leu	Ser	Ala	Pro	Pro	Thr	Pro	Ala
				155					160					165
Pro	Ile	Leu	Arg	Ala	Asp	Leu	Ala	Gly	Ile	Leu	Gly	Val	Ser	Gly
				170					175					180
Val	Leu	Leu	Phe	Gly	Cys	Val	Tyr	Leu	Leu	His	Leu	Leu	Arg	Arg
				185					190					195
His	Lys	His	Arg	Pro	Ala	Pro	Arg	Leu	Gln	Pro	Ser	Arg	Thr	Ser
				200					205					210
Pro	Gln	Ala	Pro	Arg	Ala	Arg	Ala	Trp	Ala	Pro	Ser	Gln	Ala	Ser
				215					220					225
Gln	Ala	Ala	Leu	His	Val	Pro	Tyr	Ala	Thr	Ile	Asn	Thr	Ser	Cys
				230					235					240
Arg	Pro	Ala	Thr	Leu	Asp	Thr	Ala	His	Pro	His	Gly	Gly	Pro	Ser
				245					250					255
Trp	Trp	Ala	Ser	Leu	Pro	Thr	His	Ala	Ala	His	Arg	Pro	Gln	Gly
				260					265					270
Pro	Ala	Ala	Trp	Ala	Ser	Thr	Pro	Ile	Pro	Ala	Arg	Gly	Ser	Phe
				275					280					285
Val	Ser	Val	Glu	Asn	Gly	Leu	Tyr	Ala	Gln	Ala	Gly	Glu	Arg	Pro
				290					295					300
Pro	His	Thr	Gly	Pro	Gly	Leu	Thr	Leu	Phe	Pro	Asp	Pro	Arg	Gly
				305					310					315
Pro	Arg	Ala	Met	Glu	Gly	Pro	Leu	Gly	Val	Arg				
				320					325					



<210> 12
 <211> 105
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2616827CD1

<400> 12
 Met Asn Leu Gly Val Ser Met Leu Arg Ile Leu Phe Leu Leu Asp
 1 5 10 15
 Val Gly Gly Ala Gln Val Leu Ala Thr Gly Lys Thr Pro Gly Ala
 20 25 30
 Glu Ile Asp Phe Lys Tyr Ala Leu Ile Gly Thr Ala Val Gly Val
 35 40 45
 Ala Ile Ser Ala Gly Phe Leu Ala Leu Lys Ile Cys Met Ile Arg
 50 55 60
 Arg His Leu Phe Asp Asp Asp Ser Ser Asp Leu Lys Ser Thr Pro
 65 70 75
 Gly Gly Leu Ser Asp Thr Ile Pro Leu Lys Lys Arg Ala Pro Arg
 80 85 90
 Arg Asn His Asn Phe Ser Lys Arg Asp Ala Gln Val Ile Glu Leu
 95 100 105

<210> 13
 <211> 626
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2991370CD1

<400> 13
 Met Ala Pro Ser Ala Asp Pro Gly Met Ser Arg Met Leu Pro Phe
 1 5 10 15
 Leu Leu Leu Leu Trp Phe Leu Pro Ile Thr Glu Gly Ser Gln Arg
 20 25 30
 Ala Glu Pro Met Phe Thr Ala Val Thr Asn Ser Val Leu Pro Pro
 35 40 45
 Asp Tyr Asp Ser Asn Pro Thr Gln Leu Asn Tyr Gly Val Ala Val
 50 55 60
 Thr Asp Val Asp His Asp Gly Asp Phe Glu Ile Val Val Ala Gly
 65 70 75
 Tyr Asn Gly Pro Asn Leu Val Leu Lys Tyr Asp Arg Ala Gln Lys
 80 85 90
 Arg Leu Val Asn Ile Ala Val Asp Glu Arg Ser Ser Pro Tyr Tyr
 95 100 105
 Ala Leu Arg Asp Arg Gln Gly Asn Ala Ile Gly Val Thr Ala Cys
 110 115 120
 Asp Ile Asp Gly Asp Gly Arg Glu Glu Ile Tyr Phe Leu Asn Thr
 125 130 135
 Asn Asn Ala Phe Ser Gly Val Ala Thr Tyr Thr Asp Lys Leu Phe
 140 145 150
 Lys Phe Arg Asn Asn Arg Trp Glu Asp Ile Leu Ser Asp Glu Val
 155 160 165
 Asn Val Ala Arg Gly Val Ala Ser Leu Phe Ala Gly Arg Ser Val
 170 175 180
 Ala Cys Val Asp Arg Lys Gly Ser Gly Arg Tyr Ser Ile Tyr Ile
 185 190 195
 Ala Asn Tyr Ala Tyr Gly Asn Val Gly Pro Asp Ala Leu Ile Glu
 200 205 210
 Met Asp Pro Glu Ala Ser Asp Leu Ser Arg Gly Ile Leu Ala Leu



Arg	Asp	Val	Ala	218	Ala	Glu	Ala	Gly	Val	220	Ser	Lys	Tyr	Thr	Gly	225
				230						235						240
Arg	Gly	Val	Ser	245	Val	Gly	Pro	Ile	Leu	250	Ser	Ser	Ser	Ala	Ser	255
Ile	Phe	Cys	Asp	260	Asn	Glu	Asn	Gly	Pro	265	Asn	Phe	Leu	Phe	His	270
Arg	Gly	Asp	Gly	275	Thr	Phe	Val	Asp	Ala	280	Ala	Ala	Ser	Ala	Gly	285
Asp	Asp	Pro	His	290	Gln	His	Gly	Arg	Gly	295	Val	Ala	Leu	Ala	Asp	300
Asn	Arg	Asp	Gly	305	Lys	Val	Asp	Ile	Val	310	Tyr	Gly	Asn	Trp	Asn	315
Pro	His	Arg	Leu	320	Tyr	Leu	Gln	Met	Ser	325	Thr	His	Gly	Lys	Val	330
Phe	Arg	Asp	Ile	335	Ala	Ser	Pro	Lys	Phe	340	Ser	Met	Pro	Ser	Pro	345
Arg	Thr	Val	Ile	350	Thr	Ala	Asp	Phe	Asp	355	Asn	Asp	Gln	Glu	Leu	360
Ile	Phe	Phe	Asn	365	Asn	Ile	Ala	Tyr	Arg	370	Ser	Ser	Ser	Ala	Asn	375
Leu	Phe	Arg	Val	380	Ile	Arg	Arg	Glu	His	385	Gly	Asp	Pro	Leu	Ile	390
Glu	Leu	Asn	Pro	395	Gly	Asp	Ala	Leu	Glu	400	Pro	Glu	Gly	Arg	Gly	405
Gly	Gly	Val	Val	410	Thr	Asp	Phe	Asp	Gly	415	Asp	Gly	Met	Leu	Asp	420
Ile	Leu	Ser	His	425	Gly	Glu	Ser	Met	Ala	430	Gln	Pro	Leu	Ser	Val	435
Arg	Gly	Asn	Gln	440	Gly	Phe	Asn	Asn	Asn	445	Trp	Leu	Arg	Val	Val	450
Arg	Thr	Arg	Phe	455	Gly	Ala	Phe	Ala	Arg	460	Gly	Ala	Lys	Val	Val	465
Tyr	Thr	Lys	Lys	470	Ser	Gly	Ala	His	Leu	475	Arg	Ile	Ile	Asp	Gly	480
Ser	Gly	Tyr	Leu	485	Cys	Glu	Met	Glu	Pro	490	Val	Ala	His	Phe	Gly	495
Gly	Lys	Asp	Glu	500	Ala	Ser	Ser	Val	Glu	505	Val	Thr	Trp	Pro	Asp	510
Lys	Met	Val	Ser	515	Arg	Asn	Val	Ala	Ser	520	Gly	Glu	Met	Asn	Ser	525
Leu	Glu	Ile	Leu	530	Tyr	Pro	Arg	Asp	Glu	535	Asp	Thr	Leu	Gln	Asp	540
Ala	Pro	Leu	Glu	545	Cys	Gly	Gln	Gly	Phe	550	Ser	Gln	Gln	Glu	Asn	555
His	Cys	Met	Asp	560	Thr	Asn	Glu	Cys	Ile	565	Gln	Phe	Pro	Phe	Val	570
Pro	Arg	Asp	Lys	575	Pro	Val	Cys	Val	Asn	580	Thr	Tyr	Gly	Ser	Tyr	585
Cys	Arg	Thr	Asn	590	Lys	Lys	Cys	Ser	Arg	595	Gly	Tyr	Glu	Pro	Asn	600
Asp	Gly	Thr	Ala	605	Cys	Val	Gly	Trp	Trp	610	Ser	Pro	Val	Leu	Lys	615
Val	Thr	Pro	Gln	620	Val	Gly	Lys	Ser	Leu	625	Gly	Pro				

<210> 14

<211> 296

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3031062CD1



<400> 14

Met	Glu	Trp	Trp	Ala	Ser	Ser	Pro	Leu	Arg	Leu	Trp	Leu	Leu	Leu	
1				5					10					15	
Phe	Leu	Leu	Pro	Ser	Ala	Gln	Gly	Arg	Gln	Lys	Glu	Ser	Gly	Ser	
				20					25					30	
Lys	Trp	Lys	Val	Phe	Ile	Asp	Gln	Ile	Asn	Arg	Ser	Leu	Glu	Asn	
				35					40					45	
Tyr	Glu	Pro	Cys	Ser	Ser	Gln	Asn	Cys	Ser	Cys	Tyr	His	Gly	Val	
				50					55					60	
Ile	Glu	Glu	Asp	Leu	Thr	Pro	Phe	Arg	Gly	Gly	Ile	Ser	Arg	Lys	
				65					70					75	
Met	Met	Ala	Glu	Val	Val	Arg	Arg	Lys	Leu	Gly	Thr	His	Tyr	Gln	
				80					85					90	
Ile	Thr	Lys	Asn	Arg	Leu	Tyr	Arg	Glu	Asn	Asp	Cys	Met	Phe	Pro	
				95					100					105	
Ser	Arg	Cys	Ser	Gly	Val	Glu	His	Phe	Ile	Leu	Glu	Val	Ile	Gly	
				110					115					120	
Arg	Leu	Pro	Asp	Met	Glu	Met	Val	Ile	Asn	Val	Arg	Asp	Tyr	Pro	
				125					130					135	
Gln	Val	Pro	Lys	Trp	Met	Glu	Pro	Ala	Ile	Pro	Val	Phe	Ser	Phe	
				140					145					150	
Ser	Lys	Thr	Ser	Glu	Tyr	His	Asp	Ile	Met	Tyr	Pro	Ala	Trp	Thr	
				155					160					165	
Phe	Trp	Glu	Gly	Gly	Pro	Ala	Val	Trp	Pro	Ile	Tyr	Pro	Thr	Gly	
				170					175					180	
Leu	Gly	Arg	Trp	Asp	Leu	Phe	Arg	Glu	Asp	Leu	Val	Arg	Ser	Ala	
				185					190					195	
Ala	Gln	Trp	Pro	Trp	Lys	Lys	Lys	Asn	Ser	Thr	Ala	Tyr	Phe	Arg	
				200					205					210	
Gly	Ser	Arg	Thr	Ser	Pro	Glu	Arg	Asp	Pro	Leu	Ile	Leu	Leu	Ser	
				215					220					225	
Arg	Lys	Asn	Pro	Lys	Leu	Val	Asp	Ala	Glu	Tyr	Thr	Lys	Asn	Gln	
				230					235					240	
Ala	Trp	Lys	Ser	Met	Lys	Asp	Thr	Leu	Gly	Lys	Pro	Ala	Ala	Lys	
				245					250					255	
Asp	Val	His	Leu	Val	Asp	His	Cys	Lys	Tyr	Lys	Tyr	Leu	Phe	Asn	
				260					265					270	
Phe	Arg	Gly	Val	Leu	Gln	Val	Ser	Gly	Leu	Asn	Thr	Ser	Ser	Cys	
				275					280					285	
Val	Ala	Ile	Ile	Leu	Met	Arg	Lys	Arg	Thr	Tyr					
				290					295						

<210> 15

<211> 249

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3101617CD1

<400> 15

Met	Asp	Gly	Lys	Lys	Cys	Ser	Val	Trp	Met	Phe	Leu	Pro	Leu	Val	
1				5					10					15	
Phe	Thr	Leu	Phe	Thr	Ser	Ala	Gly	Leu	Trp	Ile	Val	Tyr	Phe	Ile	
				20					25					30	
Ala	Val	Glu	Asp	Asp	Lys	Ile	Leu	Pro	Leu	Asn	Ser	Ala	Glu	Arg	
				35					40					45	
Lys	Pro	Gly	Val	Lys	His	Ala	Pro	Tyr	Ile	Ser	Ile	Ala	Gly	Asp	
				50					55					60	
Asp	Pro	Pro	Ala	Ser	Cys	Val	Phe	Ser	Gln	Val	Met	Asn	Met	Ala	
				65					70					75	
Ala	Phe	Leu	Ala	Leu	Val	Val	Ala	Val	Leu	Arg	Phe	Ile	Gln	Leu	
				80					85					90	



Lys	Pro	Lys	Val	Leu	Asn	Pro	Trp	Leu	Asn	Ile	Ser	Gly	Leu	Val
				95					100					105
Ala	Leu	Cys	Leu	Ala	Ser	Phe	Gly	Met	Thr	Leu	Leu	Gly	Asn	Phe
				110					115					120
Gln	Leu	Thr	Asn	Asp	Glu	Glu	Ile	His	Asn	Val	Gly	Thr	Ser	Leu
				125					130					135
Thr	Phe	Gly	Phe	Gly	Thr	Leu	Thr	Cys	Trp	Ile	Gln	Ala	Ala	Leu
				140					145					150
Thr	Leu	Lys	Val	Asn	Ile	Lys	Asn	Glu	Gly	Arg	Arg	Val	Gly	Ile
				155					160					165
Pro	Arg	Val	Ile	Leu	Ser	Ala	Ser	Ile	Thr	Leu	Cys	Val	Val	Leu
				170					175					180
Tyr	Phe	Ile	Leu	Met	Ala	Gln	Ser	Ile	His	Met	Tyr	Ala	Ala	Arg
				185					190					195
Val	Gln	Trp	Gly	Leu	Val	Met	Cys	Phe	Leu	Ser	Tyr	Phe	Gly	Thr
				200					205					210
Phe	Ala	Val	Glu	Phe	Arg	His	Tyr	Arg	Tyr	Glu	Ile	Val	Cys	Ser
				215					220					225
Glu	Tyr	Gln	Glu	Asn	Phe	Leu	Ser	Phe	Ser	Glu	Ser	Leu	Ser	Glu
				230					235					240
Ala	Ser	Glu	Tyr	Gln	Thr	Asp	Gln	Val						
				245										

<210> 16
 <211> 124
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3216178CD1

<400> 16
Met Gly Gly Tyr Leu Lys Thr Arg Pro Trp Thr Leu Gln His Phe
1 5 10 15
Tyr Leu Cys Leu Met Pro Ala Ala Thr Trp Leu Val Leu Leu Leu
20 25 30
Leu Leu Trp Leu Ser Leu Gly Val Lys Thr Gly Ser Cys Ser Gln
35 40 45
Pro Gln Asn Leu Cys Cys Leu Gly Thr Asp His His Cys Lys Arg
50 55 60
Gly Ser Cys Tyr Cys Asp Glu Phe Cys His Val Ala Pro Asp Cys
65 70 75
His Pro Asp His Ser Val Leu Cys Asn Pro Ala Ser Gln Met Thr
80 85 90
Lys Met Val Leu Gln Met Val Leu Arg Met Glu Asn Pro Pro Ser
95 100 105
Pro Ala Arg Ser His Leu Asp Trp Met Gln Ser Met Val Ser Ser
110 115 120
Leu Gln Val Leu

<210> 17
 <211> 101
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3406803CD1

<400> 17



```

Met Leu Pro Val Gly Ala Gln Pro Arg Ser Pro Pro Trp Val Leu
  1      5      10      15
Ala Arg Leu Leu His Pro Arg Gly Pro Ala Ala Thr Ser Leu Val
  20      25      30
Pro Phe Leu Pro Trp Gly Ser Leu Glu Ser His Thr Pro Cys Pro
  35      40      45
Tyr Arg Ala Cys Ser Pro Gly Trp Glu Leu Thr Leu Ser Thr Phe
  50      55      60
Pro Glu Arg Glu Thr Leu Ser Gly Gly Glu Val Arg Lys Arg Gly
  65      70      75
Ala Gly Ser Met Val Gly Gly Gly Glu Ser Thr Met Thr Arg Ala
  80      85      90
Leu Cys Val Arg Leu Leu Thr Lys Leu Arg Val
  95      100

```

<210> 18

<211> 540

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No. 3468066CD1

<400> 18

```

Met Ala Thr Ser Gly Ala Ala Ser Ala Glu Leu Val Ile Gly Trp
  1      5      10      15
Cys Ile Phe Gly Leu Leu Leu Ala Ile Leu Ala Phe Cys Trp
  20      25      30
Ile Tyr Val Arg Lys Tyr Gln Ser Arg Arg Glu Ser Glu Val Val
  35      40      45
Ser Thr Ile Thr Ala Ile Phe Ser Leu Ala Ile Ala Leu Ile Thr
  50      55      60
Ser Ala Leu Leu Pro Val Asp Ile Phe Leu Val Ser Tyr Met Lys
  65      70      75
Asn Gln Asn Gly Thr Phe Lys Asp Trp Ala Asn Ala Asn Val Ser
  80      85      90
Arg Gln Ile Glu Asp Thr Val Leu Tyr Gly Tyr Tyr Thr Leu Tyr
  95      100      105
Ser Val Ile Leu Phe Cys Val Phe Phe Trp Ile Pro Phe Val Tyr
  110      115      120
Phe Tyr Tyr Glu Glu Lys Asp Asp Asp Asp Thr Ser Lys Cys Thr
  125      130      135
Gln Ile Lys Thr Ala Leu Lys Tyr Thr Leu Gly Phe Val Val Ile
  140      145      150
Cys Ala Leu Leu Leu Leu Val Gly Ala Phe Val Pro Leu Asn Val
  155      160      165
Pro Asn Asn Lys Asn Ser Thr Glu Trp Glu Lys Val Lys Ser Leu
  170      175      180
Phe Glu Glu Leu Gly Ser Ser His Gly Leu Ala Ala Leu Ser Phe
  185      190      195
Ser Ile Ser Ser Leu Thr Leu Ile Gly Met Leu Ala Ala Ile Thr
  200      205      210
Tyr Thr Ala Tyr Gly Met Ser Ala Leu Pro Leu Asn Leu Ile Lys
  215      220      225
Gly Thr Arg Ser Ala Ala Tyr Glu Arg Leu Glu Asn Thr Glu Asp
  230      235      240
Ile Glu Glu Val Glu Gln His Ile Gln Thr Ile Lys Ser Lys Ser
  245      250      255
Lys Asp Gly Arg Pro Leu Pro Ala Arg Asp Lys Arg Ala Leu Lys
  260      265      270
Gln Phe Glu Glu Arg Leu Arg Thr Leu Lys Lys Arg Glu Arg His
  275      280      285
Leu Glu Phe Ile Glu Asn Ser Trp Trp Thr Lys Phe Cys Gly Ala

```



Leu Arg Pro Leu	290	Lys Ile Val Trp Gly	295	Ile Phe Phe Ile Leu	300
	305		310		315
Ala Leu Leu Phe	310	Val Ile Ser Leu Phe	315	Leu Ser Asn Leu Asp	320
	315		320		325
Ala Leu His Ser	320	Ala Gly Ile Asp Ser	325	Gly Phe Ile Ile Phe	330
	325		330		335
Ala Asn Leu Ser	330	Asn Pro Leu Asn Met	335	Leu Leu Pro Leu Leu	340
	335		340		345
Thr Val Phe Pro	340	Leu Asp Tyr Ile Leu	345	Ile Thr Ile Ile Ile	350
	345		350		355
Tyr Phe Ile Phe	350	Thr Ser Met Ala Gly	355	Ile Arg Asn Ile Gly	360
	355		360		365
Trp Phe Phe Trp	360	Ile Arg Leu Tyr Lys	365	Ile Arg Arg Gly Arg	370
	365		370		375
Arg Pro Gln Ala	370	Leu Leu Phe Leu Cys	375	Met Ile Leu Leu Leu	380
	375		380		385
Val Leu His Thr	380	Ser Tyr Met Ile Tyr	385	Ser Leu Ala Pro Gln	390
	385		390		395
Val Met Tyr Gly	390	Ser Gln Asn Tyr Leu	395	Ile Glu Thr Asn Ile	400
	395		400		405
Ser Asp Asn His	400	Lys Gly Asn Ser Thr	405	Leu Ser Val Pro Lys	410
	405		410		415
Cys Asp Ala Glu	410	Ala Pro Glu Asp Gln	415	Cys Thr Val Thr Arg	420
	415		420		425
Tyr Leu Phe Leu	420	His Lys Phe Trp Phe	425	Phe Ser Ala Ala Tyr	430
	425		430		435
Phe Gly Asn Trp	430	Ala Phe Leu Gly Val	435	Phe Leu Ile Gly Leu	440
	435		440		445
Val Ser Cys Cys	440	Lys Gly Lys Lys Ser	445	Val Ile Glu Gly Val	450
	445		450		455
Glu Asp Ser Asp	450	Ile Ser Asp Asp Glu	455	Pro Ser Val Tyr Ser	460
	455		460		465
	460		465		470
	465		470		475
	470		475		480
	475		480		485
	480		485		490
	485		490		495
	490		495		500
	495		500		505
	500		505		510
	505		510		515
	510		515		520
	515		520		525
	520		525		530
	525		530		535
	530		535		540

<210> 19
 <211> 108
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3592862CD1

<400> 19
 Met Thr Pro Ser Arg Leu Pro Trp Leu Leu Ser Trp Val Ser Ala
 1 5 10 15
 Thr Ala Trp Arg Ala Ala Arg Ser Pro Leu Leu Cys His Ser Leu
 20 25 30
 Arg Lys Thr Ser Ser Ser Gln Gly Gly Lys Ser Glu Leu Val Lys
 35 40 45
 Gln Ser Leu Lys Lys Pro Lys Leu Pro Glu Gly Arg Phe Asp Ala
 50 55 60
 Pro Glu Asp Ser His Leu Glu Lys Glu Pro Leu Glu Lys Phe Pro
 65 70 75
 Asp Asp Val Asn Pro Val Thr Lys Glu Lys Gly Gly Pro Arg Gly
 80 85 90
 Pro Glu Pro Thr Arg Tyr Gly Asp Trp Glu Arg Lys Gly Arg Cys
 95 100 105
 Ile Asp Phe

<210> 20



<211> 114
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3669422CD1

<400> 20
 Met Ser Ser Ser Ser Ser Arg Cys Leu Ser Pro Ser Pro Gly Met
 1 5 10 15
 Ser Leu Trp Ser Cys Leu Leu Phe Leu Cys Thr Pro Ser Pro Thr
 20 25 30
 Thr Thr Ser Pro Ser Pro Asp Pro Ser Gln Val Ser Thr Leu Pro
 35 40 45
 Thr Pro Ser Pro Gln Arg Glu Gly Leu Lys Gln Gly Gln Trp Arg
 50 55 60
 Lys Thr Gly Pro Ser Ser Thr His Pro His Thr Pro Ser Ser Arg
 65 70 75
 Pro Pro Ser Pro Ser Ser Leu Pro Leu Thr Trp Lys Leu Leu Gln
 80 85 90
 Pro Ile Pro Ser His Ser Leu Pro His Pro Pro Lys Ile His Thr
 95 100 105
 Gly Pro Ser Leu Ala Glu Cys Gly His
 110

<210> 21
 <211> 114
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3688740CD1

<400> 21
 Met Arg Gly Glu His Asn Ser Thr Ser Tyr Asp Ser Ala Val Ile
 1 5 10 15
 Tyr Arg Gly Phe Trp Ala Val Leu Met Leu Leu Gly Val Val Ala
 20 25 30
 Val Val Ile Ala Ser Phe Leu Ile Ile Cys Ala Ala Pro Phe Ala
 35 40 45
 Ser His Phe Leu Tyr Lys Ala Gly Gly Ser Tyr Ile Ala Ala
 50 55 60
 Asp Gly Ile Ser Ser Leu Cys Tyr Ser Ser Leu Ser Lys Ser Leu
 65 70 75
 Leu Ser Gln Pro Leu Arg Glu Thr Ser Ser Ala Ile Asn Asp Ile
 80 85 90
 Ser Leu Leu Gln Ala Leu Met Pro Leu Leu Gly Trp Thr Ser His
 95 100 105
 Trp Thr Cys Ile Thr Val Gly Leu Tyr
 110

<210> 22
 <211> 287
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3742589CD1



<400> 22
Met Glu Leu Glu Arg Ile Val Ser Ala Ala Leu Leu Ala Phe Val
1 5 10 15
Gln Thr His Leu Pro Glu Ala Asp Leu Ser Gly Leu Asp Glu Val
20 25 30
Ile Phe Ser Tyr Val Leu Gly Val Leu Glu Asp Leu Gly Pro Ser
35 40 45
Gly Pro Ser Glu Glu Asn Phe Asp Met Glu Ala Phe Thr Glu Met
50 55 60
Met Glu Ala Tyr Val Pro Gly Phe Ala His Ile Pro Arg Gly Thr
65 70 75
Ile Gly Asp Met Met Gln Lys Leu Ser Gly Gln Leu Ser Asp Ala
80 85 90
Arg Asn Lys Glu Asn Leu Gln Pro Gln Ser Ser Gly Val Gln Gly
95 100 105
Gln Val Pro Ile Ser Pro Glu Pro Leu Gln Arg Pro Glu Met Leu
110 115 120
Lys Glu Glu Thr Arg Ser Ser Ala Ala Ala Ala Asp Thr Gln
125 130 135
Asp Glu Ala Thr Gly Ala Glu Glu Glu Leu Leu Pro Gly Val Asp
140 145 150
Val Leu Leu Glu Val Phe Pro Thr Cys Ser Val Glu Gln Ala Gln
155 160 165
Trp Val Leu Ala Lys Ala Arg Gly Asp Leu Glu Glu Ala Val Gln
170 175 180
Met Leu Val Glu Gly Lys Glu Glu Gly Pro Ala Ala Trp Glu Gly
185 190 195
Pro Asn Gln Asp Leu Pro Arg Arg Leu Arg Gly Pro Gln Lys Asp
200 205 210
Glu Leu Lys Ser Phe Ile Leu Gln Lys Tyr Met Met Val Asp Ser
215 220 225
Ala Glu Asp Gln Lys Ile His Arg Pro Met Ala Pro Lys Glu Ala
230 235 240
Pro Lys Lys Leu Ile Arg Tyr Ile Asp Asn Gln Val Val Ser Thr
245 250 255
Lys Gly Glu Arg Phe Lys Asp Val Arg Asn Pro Glu Ala Glu Glu
260 265 270
Met Lys Ala Thr Tyr Ile Asn Leu Lys Pro Ala Arg Lys Tyr Arg
275 280 285
Phe His

<210> 23
<211> 854
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 078811CB1

<400> 23
atcttgccctc gtggacccaa aggtagcaat ttgaaacatg aggagtaaga ttctactgtt 60
ttgtcttcta ggatcaactc ggtcattacc agtctttcct tctttaagtc tgataccatt 120
aacacagatg ctacacactg ggccagatct gcatctgtta aatcctgctg caggaatgac 180
acctgggtacc cagacccacc cattgacct gggaggggtg aatgtacaac agcaactgca 240
cccacatgtg ttaccaatct ttgtcacaca acttggagcc ccagggcact atcctaagct 300
cagaggaatt gccacaaatc ttacagagcc tcatcatcca ttcttgttc cccgggaggc 360
atccttgccc accagtcagg caggggctaa tccagatgtc caggatggaa gccttccagc 420
aggaggagca ggtgtaaatc ctgccaccca gggaacccca gcaggccgcc tcccaactcc 480
cagtggcaca gatgacgact ttgcagtgc cacccttgca ggcattccaa ggagcacaca 540
tgccatcgag gaagccacca cagaatcagc aaatggaatt cagtaagctg ttccaaatct 600
tttcaactaa gctgcctcga atttgggtgat acatgtgaat ctttatcatt gattatatta 660
tggaatagat tgagacacat tggatagctc tagaagaaat taattcttaa tttacctgaa 720
aatattcttg aaatttcaga aaatatgttc tatgtagaga atcccaactt ttaaaaaacaa 780



taattcaatg gataaatctg tctttgaaat ataacattat gctgectgga tgatatgcat 840
attaaaaacga atta 854

<210> 24
<211> 1804
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 371156CB1

<400> 24
gtgataggca gctttccttc ttttcaacag tgatacctac gaaaatcaaa ataaatgcaa 60
gttgaggttt tgtgctcaat gaaagggctg tcaaccccag aaggccgaca caaaaaaaat 120
ggtatgtgaa gatgcaccgt cttttcaaat ggccctgggag agtcaaattg cctgggagag 180
ggggccctgcc cttctctgct gtgtccttct ggcttcccag ttgagctccc aagaccagga 240
cccactgggg catataaaat ctctgctgta tcctttccggc tcccagttg agctcccaag 300
accaggaccc actggggcat ataaaaaagt caaaaatcaa aatcaaacaa caagttctga 360
gttacttagg aaacagactt cgcatttcaa tcagagaggc cacagagcaa ggtctaaact 420
tctggcttct agacaaatct ctgatagaac atttaaatgt gggaagtggc ttcccaggt 480
cccacccctt gtttagggat agagttgata tcatttttat aggtgccatg tatgectctg 540
cctgaatttt ttttaattgac ttttgagctt ttgagattgc acgagggaga acaaggcctt 600
tctctgttgt gataggaaag acttaacctt aaattaaacc agcaagaaag cattagtaaa 660
auctaatacaa tatgaagggc tcttatgagt catttttttc aaaagatgaa aactccagaa 720
acgcacagga acgaaatacc tcccagaaac atgaagcaat catcgaagac tcactggtaa 780
tattttttaa aagtatacag atcaaagcaa aaagaagcca tgtgtaacaa agagaaatgt 840
gcaaatattt ttttaaggcag tattaagtgc aagaggagta acatgaaata aacattcttt 900
cacatggcta ctgggaatat aaatttctgt ccagaaagge cgtagcagtt tgacgatagg 960
tggcaaaaacc ttaagattgt gtactggggc ccagaatttt tatttctagg aatgtatcct 1020
gaggaaatta tccgagatcc ccacaaactg caatgttttag gaattgtcct tatagcattg 1080
catacacaag aaaaacagag aaaagccctg tcctgtcag tggaaaaggg gttcaatgaa 1140
ttacgggtgt tctgcatgag gcttttatga cattaaaaat tgttgaacaa cggccaggca 1200
cagtggctca tgccctgtaat cctaacactt tgggaggcca aggtgggaag attgectgag 1260
ctcaggagtt tgagaccagc ctgggcaaca cggtgaaacc cgtctctac taaaatacaa 1320
aaaattagcc gggcgtcgca gcatgcgct gtagtcccag ctgctcagga ggctgaggca 1380
ggagaattga ttgaaccggg gaggcagagg ttgcaactg ctgagattaa gccaccgcac 1440
tccagcctgg gcgacagagc aagattccgt tcccagaaa aaaaaattgt tcaacaataa 1500
gggcaaaagg agagaatcat aacatctgat taaacagaaa aagcaagatt tttaaaacta 1560
actatataag gatgggtcca gctgtgtcaa aaggaaagct gtttgtaata cgtgtgcata 1620
aaaattaaat agaggtgaac acaattattt taaggcagtt aaattatctc tgtattgtga 1680
actaagactt tctagaattt tacttattca ttctgtactt aaattttttc taatgaacac 1740
atatactttt gtaatcagaa aatattaaat gcatgtattt ttcaaaatca aaaaaaaaaa 1800
aaaa 1904

<210> 25
<211> 2663
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 584050CB1

<400> 25
ggagaaaagga tggccggcct ggccggcgcg ttgggtcctgc tagctggggc agcggcgctg 60
gcgagcggct cccagggcga ccgtgagccg gtgtaccgag actgcgtact gcagtgcgaa 120
gagcagaact gctctggggg cgtcttgaat cacttccgct cccgccagcc aatctacatg 180
agtcttagcag cgtggacctg tccggacgac tghtaagtat agtgatgtg ggtcacccgt 240
gggctctacc tccaggaagg tcacaaagtg cctcagttcc atggcaagtg gcccttctcc 300
cggttcctgt tctttcaaga gccggcatcg gccgtggcct cgtttctcaa tggccctggc 360
agcctgggtga tgctctgccg ctaccgcacc ttcgtgccag cctcctcccc catgtaccac 420
acctgtgtgg ccttcgcctg ggtgtccctc aatgcattgt tctggteccac agtcttccac 480
accagggcaca cttgacctac agagaaaaat gactactct gtgcctccac tgtcatccta 540
cactcaatct acctgtgctg cgtcaggacc gtggggctgc agcaccagc tgtgggtcagt 600



```

gcttccggg ctctctctgt gctcatgctg accgtgcaag tctctlaact gagctcato 660
cgcttcgaat atggctacaa cctgggtggc aacgtggcta ttggcctggt caacgtggtg 720
tggtggctgg cctgggtgct gtggaaccag cggcggtgtg ctcaagtgag caagtgcgtg 780
gtgggtggtt tgctgctgca ggggctgtcc ctgctgagag tgcttgactt cccacgcctc 840
ttctgggtcc tggatgacca tggcatctgg cagatcagca ccatacctgt ccacgtcctc 900
tttttcagct ttctggaaga tgacagcctg tacctgtga aggaatcaga ggacaagtcc 960
aagctggact gaagaccttg gagcgagtct gccccagtgg ggatcctgac cccgcccctg 1020
tggcctccct tctccctcca acccttgaga tgattttctc ttttcaactt cctgaacttg 1080
gacatgaagg atgtgggccc agaactcatg ggccagacca cccctgttg gccctcacca 1140
gccttggagt ctgttctagg gaaggcctcc cagcatctgg gactcgagag tgggcagccc 1200
ctctacctcc tggagctgaa ctggggtgga actgagtgtg ttcttagctc taacgggagg 1260
apagctgctt gtttctctcc caccagcctc ctccccacat cccagctgc ctggtggtg 1320
cctgaagccc tctgtctacc tgggagacca gggaccacag gccttaggga tacagggggt 1380
ccccttctgt taccaccccc caccctcctc caggacacca ctagggtgtg ctggtgctt 1440
gttcttttgg cagccaaggt tcacggcgat tctccccatg ggatcttgag ggaccaagct 1500
gctggggttg ggaaggagtt tcacctgac cgttgcccta gccaggctcc caggaggctc 1560
caccatactc cctttcaggg ccagggtctc agcaagacca gggcaaggat cctgtgctgc 1620
tgtctgggtg agagcctgcc accgtgtgtc gggagtgtgg gccaggctga gtgcataggt 1680
gacagggcgg tgagcatggg cctgggtgtg tgtgagctca ggcctagggt cgcagtgtgg 1740
agaagggtgt tgtcggggaa gaggtgtggc ttcaaagtgt gtgtgtgcag ggggtgggtg 1800
tggttagcgt gggttagggg acgtgtgtgc gcgtgctggg gggcatgtga gatgagtga 1860
tgccgggtga ttgtgtccca gttgagaggt tggagcagga tgagggaatc ctgtcaccat 1920
caataatcac ttgtggagcg ccagctctgc ccaagacgcc acctggggcg acagccagga 1980
gctctccatg gccaggctgc ctgtgtgcac gttccctgtc tggtgcccc tggcccgcct 2040
cctgcaaac tcacagggc cccacacaac agtgccctcc agaagcagcc cctcggagge 2100
agaggaagga aaatggggat ggctggggct ctctccatcc tcttttctc cttgcccctc 2160
catggctggc ctccccctcc aaaacctcca tccccctgt gccagcccc ttgccatagc 2220
ctgatttttg ggaggaggaa ggggcgattt gagggagaag gggagaaagc ttatggctgg 2280
gtctgggttc ttcccttccc agagggtctt actgttccag ggtggcccca gggcaggcag 2340
ggccacact atgctgcgc cctggtaaa ggtgacccct ccatttacca gcagccctgg 2400
catgttctct cccacagga atagaatgga gggagctcca gaaacttcc atcccaaagg 2460
cagtctcctg ggttgaagca gactggattt ttgctctgcc cctgacccct tgccctctt 2520
tgagggaggg gagctatgct aggactccaa cctcagggac tcgggtggcc tgcgctagct 2580
tcttttgata ctgaaaactt ttaaggtggg agggtggcaa gggatgtgct taataaatca 2640
attccaagcc tcaaaaaaaa aaa 2663

```

<210> 26
 <211> 769
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 863808CB1

```

<400> 26
gcgacgccga cgcaaggctg ctgctatggg gccggggcgc cgtgtggcgc ggctgctcgc 60
cccactaatg tggcgcaggg cggtttctct ggtggcgggg tccgcggttg gagccgagcc 120
cgggcttcgt ctgctggccg tgcagcggct tcccgtagga gcagcgttct gccgggcttg 180
ccagacccca aactttgtcc gcggcctgca cagcgagcct gggctggagg agcgggcgga 240
ggggacggtc aacgagggac gcccagaatc ggacgcggca gatcatactg gtcccaagtt 300
tgacatcgat atgatgggtt cacttctgag gcaagaaaat gcaagagaca tttgtgtgat 360
ccaggttcct ccagaaatga gatatacaga ttactttgtg attgttagtg gaacttctac 420
ccgacactta catgccatgg ccttctacgt tgtgaaaatg tacaaacacc tgaaatgtaa 480
acgtgacctt catgttaaga tagaaggga ggacactgat gactggctgt gcgtggattt 540
tggcagcatg gtgattcatt tgatgcttcc agaaaccaga gaaatctatg aattagagaa 600
attatggacc ctacgttctt atgatgacca gttagctcag atagcacctg agacagtacc 660
tgaagacttc attcttggaa tagaagatga tacttcactt gtgactccag tggaggtaaa 720
atgtgaataa aatattttat gcactgcgtt agtcaaaaaa aaaaaaaaaa 769

```

<210> 27
 <211> 1257
 <212> DNA
 <213> Homo sapiens



<220>
 <221> misc_feature
 <223> Incyte ID No: 978433CB1

<400> 27
 gagggcggcg ggtgaaaggg gcattgcatgc agcctggcggc ggccctgggag cggggcgggag 60
 cagacggtga ccacgttccct ctccctgggct tccctccgctt ccagctccggt gctgcccgggc 120
 acctggggagc catgcgacccc caggggccccg ccgctctcccc gcagcgggctc cggggcctccc 180
 tgcctgctcct gctgctgcag ctgcccgcgc cgtcgagcgc ctctgagatc cccaagggga 240
 agcaaaaaggc gcagctccgg cagagggaggc tgggtggacct gtataatgga atgtgcttac 300
 aagggccagc aggagtgccct ggtcgagacg ggagccctgg ggccaatggc attccgggta 360
 cacttgggat cccaggtcgg gatggattca aaggagaaaa gggggaatgt ctgagggaaa 420
 gttttgagga gtccctggaca cccaactaca agcagtggtc atggagttca ttgaattatg 480
 gcatagatct tgggaaaaatt gcggagtgtc catctacaaa gatgcgttca aatagtgctc 540
 taagagtttt gttcagtggtc tcaattcggc taaaatgcag aaatgcctgc tgtcagcgtt 600
 ggtatttcaac attcaatgga gctgaatggt caggacctct tcccattgaa gctataattt 660
 atttggaacca aggaagccct gaaatgaatt caacaattaa tattcatcgc acttctcttg 720
 tgggaaggact ttgtgaagga attgggtgctg gattagtgga tgttgctatc tgggttggca 780
 ctgtttcaga ttacccaaaaa ggagatgctt ctactggatg gaattcagtt tctgcgcatc 840
 ttattgaaga actacaaaaa taaatgcttt aattttcatt tgcctacctc ttttttatta 900
 tgccttggaa tgggtcactt aaatgacatt ttaataaagt ttatgtatac atctgaatga 960
 aaagcaaaagc taaatatggt tacagaccaa agtgctgatt cacactgttt taaaatctag 1020
 cactattcat tttgcttcaa tcaaaagtgg ttccaatatt ttttttagtt ggtagaata 1080
 cttctctcat agtcacattc tctcaacctc taatttggaa tattgtgtg gtcttttgtt 1140
 tttctcttta gtatagcatt tttaaaaaaa tataaaagct accaatcttt gtacaatttg 1200
 taaatgttaa gaattttttt tatatctggt aaaaaaaat tatttccaac aacctta 1260

<210> 28
 <211> 2560
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1655369CB1

<400> 28
 ttccagtga gagcaagtgc tgcccgaccc aggacctgt gccaggctag cagccctcca 60
 gctccctcca gagaggaaac ctctgtcttg ctgaggggtg gactagctgg gatgtctcac 120
 tccagttgct caggttcacc caggatggag ctccgtggag tggccagcct gattctagcc 180
 ctgtctcttc tggcagcaca tgccacacct gcctgggctt tctgtccct gatgtctac 240
 gagccctgc ctctcaatg ttctcaaaag acagaccccc ctgaggccag cttgaatgtg 300
 aagactgctg aagtcagctg gcttcaactt agctgcagaa aaggtggctg ggatggccca 360
 ggtgcaccca gagggcccag ccttttggct gcctttgggt tgtgacttgg gttgtctctg 420
 aggccttgc agagctgggc ctgcccgttg tggcggtcc gacctcgggc agtcagtgct 480
 ccgcagcctc agcactgcat cccagaccca gtgtctcag aggggaagag cagcctccc 540
 gcctcatgga accaggagtc ccaaaaagtc aggagcctgg aggtctctgaa aggagcaggg 600
 attccatagt gcgtgaagct gaaataggcg cctcctggg gagccccag caaaactgtt 660
 ttccataccc actcccagaa ctgccccgct ccagctccag cgcagcgcgc agctgggttg 720
 caggcgtcat tggagaggcc tggctgcccc aggggcagca gggagtgggt gacctgtatg 780
 ggctggcagg aggcatttgg ccattgctgac aagtgtcacc tgccttcta gctggagcc 840
 accctcagg tggcctgctt gcacctccta tccggaggta gcctgcccc cctgtaggca 900
 gagggggctc ttgcttgagg cctgcacagg aagcaagtat agccccgggt cccagagtg 960
 ggttccactt agccctgggc agatggcctg tctgagatc tctgtccca gacccccca 1020
 tctggggagc acagtcctta ggctgcccgt tccaggaagg ggggtgcggct ctgtcaggaa 1080
 acctggactc tcaaggccca ccagcctctc cgtgagtgtt agaaatcaca gatacagtat 1140
 atacttaatt acactaaatt attgctggga ttctttataa gactaatta tacctgatta 1200
 taggttaaaa tattttatctt gtcaaaaatat ttcttggga atgtgtttta cctttctgc 1260
 gttcattggt gctgagatgt gaaaactaac cattcccccc tgcctacctt tttggccact 1320
 gggcggcgca gaatggcgt atgtgcagtt gggcccttgg caccatgggc ccttggcctg 1380
 cctgctgcag agtagccctg cctgggcagt ctccaggcac tgagcaggcc atctgtggcc 1440
 aggctgagag aatgactggc tgccttacca gcgtgcattg gacaaggagc tttggagcct 1500
 caagggggtt ttgctggcct gggctagagg gaaaggtgac catcctctct tccctctgtc 1560
 tttctattag cgcctccatg tgaagtgttg tgccttgggt cactagcctt cccccaccac 1620
 cccaccatgc cacctgggtg tcttggggcc tctgtctca ctccagcccc tggggaggag 1680
 aggaccagc ccggagagtt ggggcaaggc ctccacatgc cccaagggca acagatgctc 1740




```

gcaggggcagg tggctggcagat gctcacgctc ctggccccct ccttcccgcg gccacacccc 1860
accttggggc cccgcagaca cgcctctctc actcagttgg gccacagcct ctggatggct 1865
tggggtaggg catggggcaca cctgggggca gggcagcccc tggggcagct ctggaagagt 1870
agtgtggagg agcacttggc tgcagcctgg ctccagcctc tgggaactgt ggagtgggtc 1875
ctgggagctt ctgcaactgtc ggcttttggg acgtctcacc cacttggggt acagtaggcc 1880
ttccccccc agagagaagt gttttccccc cagagacatt gttgttcaga ccttgaagtg 1885
ctggcctccc ccagtggccc tcaccagccc ttcttatctg tgggggtccaa ctccaggcttc 1890
ccctggggcc accagccata gggagcagcc atcaggcccc gagtccagaac tgcctctgtc 1895
tctccataac ccaggctctt cccggagagg gggacggata tttatttctt aaagtttgca 1900
cttaattgtg aggattctca ggattgttgg gggtactga aaagagggaat gtgttgaatg 1905
ctggctttgg tctccattgg tcttagaagt ttagtgtttt tctcactgtc atgtgtttct 1910
gtgggcagag ctggttttgg aggttgggtc agtggacccc aggtccagag catccatcca 1915
ccccactggc cctccttcca gataccctct ctctaattgg gttcttgcac gtaaaatact 1920
ccacaataaa taaataattg aacaaattaa aaaaaaaaaa 1560

```

<210> 29
 <211> 614
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1703244CB1

```

<400> 29
gtgcaagagg aacaaagaaa gggactcctg cagcgtcccg ctgacctggc ccttgtcata 60
tatctcctcc ttgctggctt cttcactctg ttccgggggc tgggtgggtg tgattgccc 120
acagatgcct gctttgtcta tatctaccag tatgagccat acctgcccga ccttgtggcc 180
taccctaagg tgcagatgct gatgtacatg ttttatgtcc tgcctttctg cggcctggct 240
gcctatgtcc tcacctccc ttggttgcctc ttgcttccag actgggcctt ggtgtttgct 300
ggaggtcctg gccaggcaca gtctctgcac atgggggctt ccattgcacc gcgcacccc 360
ttcacctacc gtgtgcttga ggacacctgg ggtgtcttct tctgttgcaa tctgctgtat 420
gcgctggggc cccacctgct ggctaccgt tgccttcagt ggcccgcatt cttccaccag 480
ccaccacctt ccgacccctt agcctccac aagaagcagc attgagagag ctgtggactc 540
aggaccagc actctgttta cgtgcccagt cagccctacc tggggaagcg ggggttgggt 600
gttttagaga cagg 614

```

<210> 30
 <211> 1936
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1730819CB1

```

<400> 30
gactacgggc tcacagccgt ccttctcctg gtgggaagaa gccgagatgg cggcagccag 60
cgctggggca acccggtgc tcttgccttt gctgatggcg gtagcagcgc ccagtcgagt 120
ccggggcagc ggctggccgg ccgggaactgg tgcggagggg gctggggcgg aaggtcgaga 180
gggcgaggcc tgtggcacgg tggggctgct gctggagcac tcatttgaga tggatgacag 240
tgccaacttc ccgaagcggg gctcactgct ctggaaccag caggatggta ccttgtccct 300
gtcacagcgg cagctcagcg aggaggagcg gggccgaact cgggatgtgg cagccctgaa 360
tggcctgtac cgggtccgga tcccaaggcg acccggggct ctggatggcc tggaaagtgg 420
tggctatgtc tctccttttg tccctgcctg ctccctgggt gagtgcgacc tgtcggaaca 480
gctgaccttg cactgggatg tggccggcaa cgtggtgggc gtgtcggtgg tgacgcacc 540
cgggggctgc cggggccatg aggtggagga cgtggacctg gagctgttca acacctcgg 600
gcagctgcag ccgcccacca cagcccagc ccttgagacg ggggccttca ttgagcgct 660
ggagatggaa caggcccaga aggccaaaga ccccaggag cagaagtcct tcttcgcaaa 720
atactggatg tacatcattc ccgtcgtcct gttctctatg atgtcaggag cgcagacac 780
cgggggcccag ggtgggggtg ggggttgttg tgggtgggtgg ggtagtggcc ggtgagggcc 840
caggctggct agcgtcccgt cttgcacacc caggggcctc cctttctgct ggagtccct 900
gtgtcctcag ccattcccaag aagggtttgc tggctccctc tttccccccg tcccacgagg 960
ccacctgggc cagccccctg tctctgctct tctgctggca gaggagcagc tggactgggg 1020
cctttggcac agcagccggt gtctcctgct ccgcctccc ccattggccc atgcagcccc 1080

```



```
agggggttcc cccctgcccc tggagtagag cccagatcc tggccactat gccagttctg 1140
acctggcacc cccctaccoc gagcccatgc agtctgggaa catgccgcct tctctccagg 1200
ctctgtgccc ttgttccagg tggctccacc ctctgtccc tggctggggt aggtgggtcc 1260
gtccaggccc ctgcagcgcc cccctcaatt tgagactgga ctaggatgca ggcctccctc 1320
tgtgtccccc tgaggggtacc ctgggtccccc tcatcagggg cagaggcatg aaagagtgcg 1380
ggctgggaggt cgggggggctt ctgggcccga cgcctagtga agcccccggg gtctgtggtt 1440
gacatttgtc tgcctgggtgc aaacaaggaa tctttgctt taaggtgaca ggccctccac 1500
aggttccagc atttgaaggga aaaggtttta gaaagaaaaa aaaaccaaca gttagtggag 1560
tcaaaagccc gacactgtaa atagaacccc ctccaccacc ccccgccgcb cagcatccca 1620
cttgagactg ggtgtctacga ggccctggcg gcttttgcct tgtgccaccc tccctgttaag 1680
tctatttaaa aacatcgacg atacattgaa atgtgtgaac gttttgaaaa gttacagctt 1740
ccaccacccc aaagcaactg ttgttttggc aagacgggtc tgatgtacaa ccttgattga 1800
aattccactg tcaacttgata cgttatccag aaacccaagg aatggctgtc cccatcccca 1860
tgttggctgt tggagctcag ctgtgttgtg tggcagttta ttaactgtc cccagatcg 1920
acaggcaaaa aaaaaa 1958
```

<210> 31
<211> 1958
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Inocyte ID No: 1757161CB1

```
<400> 31
gcgcgcgcct cagctacggc ccgagcgagc ccgcgcgcgc cgggcccggc cacagcctgc 60
agcggagccc acgagaggca gcgcccattg ggagcagacc tactcgtggg cctattccct 120
ggtggattcc agtcaagtgt ctacatttct gatttccatt cttcttatag tctatggtag 180
ttccaggctc cttaatatgg actttgaaaa tcaagataag gagaaagaca gtaatatgtc 240
ttctgggtct ttcaatggca acagcaccaa taatagcatc caaacaattg actctaccca 300
ggctctgttc cttccaattg gagcatctgt ctctctttta gtaatgttct tctcttttga 360
ctcagttcaa gtagttttta caatatgtac agcagttctt gcaacgatag cttttgcttt 420
tcttctcttc ccgatgtgcc agtattiaac aagaccctgc tcacctcaga acaagatttc 480
ctttgggtgc tgtggacgtt tcaactgctg tgagttgctg tcattctctc tgttctgcac 540
gctctgcttc atctgggttc tcaactggcc ttggcttctc atggatgcac tggccatggg 600
cctctgtgtc gccatgatcg cctttgtccg cctgcgcgag ctcaaggctc cctgcctgct 660
tctctcaggg cttctcatct atgatgtct ttgggtattt ttctcagcct acatcttcaa 720
tagcaacgtc atggtgaagg tggccactca gcgcgctgac aatccccctg acgttctatc 780
ccggaagccc cacttggggc ccaatgttgg gcgtgatgtt cctcgccctg ctctgcctgg 840
aaaactggtc ttccaagct ccactggcag ccacttctcc atgttgggca tgggagacat 900
cgttatgccc ggtctcctac tatgttttgt ccttcgctat gacaactaca aaaagcaagg 960
cagtggggag tctgtggggg cccctggacc tgccaacatc tccgggcgca tgcagaaggt 1020
ctcctacttt cactgcaccc tcatcggtat cctttgtagg ctgctcactg ctactgtggg 1080
gtctcgcaat cacggggccg cccagcccg cctctctat ttgggtgccat ttactttatt 1140
gccactcttc acgatggcct atttaaaggg cgaactccgg cggatgtggg ctgagccttt 1200
ccactccaa g tccagcagct cccgattcct ggaagtatga tggatcacgt ggaaagtga 1260
cagatggccg tcatagtccc tttctctcaa ctcatgggtt gtttccctct agagctggcc 1320
tggtactcag aaatgtacct gtgttttaagg aactgcctgt tgactggatt tggcatttaa 1380
agggagctcg ttgtcaggag agaggtgctg gagccctgt ttggttcctt tcttccctgc 1440
gatgtagagg tggggccccc tccaagaggg acaggcctct cccagcgcg gagatggaga tttgactgtt 1500
acgtttttat ggatctgcac cagactgtta ccttctgggg gagatggaga tttgactgtt 1560
taaaaactga aaacagcgag gagtctttct agaacttttg aacactaaaa ggatgaaaaa 1620
aattagcaaa ccgaagtttc ttcaatgacc cctcgagaa cttgggacca gtttccctatg 1680
ggggactcag ttctagagaa ctgagacaga agctctctct tcgttatatt cttctttcct 1740
tttttggat ttttaataa ttttctgtgg tgtgaagtga cttattaaat ccacagacat 1800
tgagtgaact cttacaacat ccacataaga atttgttgta atgagttcat gtccacccag 1860
atgttgtgtt ggcagtgaac aagggcacgg tttttatata tacgtacata tatatatata 1920
aacacacaca tagatatata tgaataaaca aaaatgat 1958
```

<210> 32
<211> 1424
<212> DNA
<213> Homo sapiens



<220>
 <221> misc_feature
 <223> Incyte ID No: 1976095CB1

<400> 32
 gtcaagtagg agacaaggag caaagtcccta tcacagcggg aggggacgca agggcctgca 60
 gaggctgagg agggaaaaaag ccagtgcgcc agcgggaagca cagctcagag ctggtctgca 120
 atggacatcc tgggtcccact cctgcagctg ctggtgctgc ttcttacctt gcccctgca 180
 ctcatggctc tgcctgggctg ctggcagccc ctgtgcaaaa gctacttccc ctacctgatg 240
 ggcgtgctga ctcccaagag caacccgcaag atggagagca agaaaaggga gctcttcaga 300
 cagataaagg ggcttacagg agcttcgggg aaagtggccc tactggagct gggctgggga 360
 aaggagagca actttcagtt ctacccaacc ggctgcaggg tcacctgccc agacccaaat 420
 ccccactttg agaagtctct gacaaagagg atggctgaga acaggcacct ccaatatgag 480
 cgggtttgtg tggctcctgg agaggacatg agacagctgg ctgatggctc catggatgtg 540
 gtggctctgca ctctgggtgt gtgctctgtg cagagcccaa ggaaggctct gcaggaggtg 600
 cggagagtag tgagaccggg aggtgtgctc tttttctggg agcatgtggg agaaccatat 660
 ggaagctggg ccttcagtgt gcagcaagtt ttctagcccc cctggaaaaca cattggggat 720
 ggtgtgctgc tcaccagaga gacttggaag gatcttgaga atggccagtt ctccgaaatc 780
 caaatggaa gacagccccc tcccttgaa tggctacctg ttggggcccc catcatggga 840
 aaggctgtca aataatcttt cccaagctcc aaggcactca ttgtctctt cccagacctc 900
 caattagaa aagccaccca ccagcctatc tatcttccac tgagagggac ctagcagaat 960
 gagagaagaa attcatgtac caactcctag tccctctctc ccaaacctct gccagggcaa 1020
 tctctaaact caatcccgcc ttgacagtg aaaaagctct actttctacg tgacccaggg 1080
 aggaacact aggaacctgt tgtatctcca actgcaagtt ttctggactg tctcccaacg 1140
 tttgcctccc aatgtttgtc ctttctcttg ttcccatggt aaagctctc tcgctttcct 1200
 cctgaggcta caccatggg tctctaggaa ctggtcacaa aagtcattgt gcttgcatc 1260
 ctgccaaagg cccctgacc tctctcccca ctaccacctt cttcttgagc tgggggcacc 1320
 agggagaatc agagatgctg gggatgccag agcaagactc aaagaggcag aggttttgtt 1380
 ctcaaatatt ttttaataaa tagacgaaac caccgaaaaa aaaa 1440

<210> 33
 <211> 2238
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2169991CB1

<400> 33
 cctgtctgtg cacacctcac gggaaggggc agcctgtttc ctcccggtca cctccaaatc 60
 ttgtctgctt taattcaact cagaggtatg cacttgaggt agggaggcag ggggaagtgg 120
 gatggcagga catggatggc ccttgaggca ttggctctgg gtgtcatggg ctgtgagagt 180
 caagaagggg agtggcctgc ctgacttggg ttcgaaaggg tcactctggc cactgagggt 240
 agaactgaaa ctagccaggc caattctcca ttgttccctg tcttccaggt aggagaatat 300
 tagattaggt ctggagtcag aaggtagcag gggctggggg ttgcaggggg atgttgagaa 360
 gaagtggctc ttgggtcagg gtgggaagcc aacaggattt cctggtgcat tggagggtgaa 420
 agggaaagagg cgtctgttgt tgtgctttgg gcttgagcag ccagaagccg ttgccatcac 480
 ctatgacaag ggagacagcg cctggggcag agcccagtat ggggtgcatt cagggttagat 540
 caggaggtat gagtgggcag tgggogagcc aggttggggg ttgtggggga ggcctggggt 600
 gcaggtttaa gcgtgggacc cactgcagat tgggtgtggg tcattgcatt tggggtggct 660
 ctcaggctct cctgctggc tctccactcc caggggcttt ctctcccgag attccttagc 720
 tgggtgaggg gcaggacaga gccctttcct agggaaagccc ggcacccctt gctgtccagg 780
 gaaggggagt ccttctagcc cctgacagct tctctgcccc tcccttggcc tccccaggcc 840
 tagccagggt tgagttctca cccactgtg ccgcctgccc ttgttacctg gaagcacagc 900
 cttggggact gaggaggccc tcaactgtca ttaagaagg gaatcagcca atttgtgctc 960
 accacctctg gggaaggtgt gagaggagag aaggaagtgg ctgtttggct gctgacaaca 1020
 tgaagacttc ctgcgatgag aacagaggca caggtgcccg cctgagccc cctgggtgct gctgaccttg 1080
 ggaactggagg gggccatggg gcaccggacc ctggctcctg cctgggtgct gctgaccttg 1140
 tgtgtcactg cggggacccc ggaggtgtgg gttcaagttc ggatggaggg caccgagctc 1200
 tcttcttcca ccatcgttg tgggttcttg ggtctgtgct ccatctccct ggtgactgtg 1260
 agctgggggg gccccaacgg tgcctggggg accacgctgg ctgtgttgca cccagaacgt 1320
 ggcacccggc aatgggcccc tgctcgccag gcccgctggg aaaccagag cagcatctct 1380
 ctcatcctgg aaggctctgg ggccagcagc cctgagccc acaccacct ctgctgcaag 1440
 tttgctcct tccctgaggg cctctgggag gctgtggga gctcccgcc cagctcagac 1500
 ccagggtctt ctgccccgcc gactcctgcc cccattctgc gggcagacct ggcggggatc 1560



```

ttggggggtct caggagtcct cctcttttggc tgtgttctaac tctttcctct gctggcgccga 1620
cataagcacc gccctgcacc taggctccag cgttcaccda ccagcaccda ggcaccgaga 1680
gcacgagcat ggjcaaccaag ccaggcctcc caggctgctc ttccagtcce ttatgcact 1740
atcaacacca gctgcgcgcc agctactttg gacacagctc accccctagg ggggcgctc 1800
tgggtggggt cactccccc ccaagctgca caccggcccc agggccttga cgcctgggc 1860
tccacaccca tccctgcaag tggcagcttt gtctctgttg agaattgact ctacgctcag 1920
gcaggggaga ggccctctca cactggctcc ggctcactc tttctcttga cctcggggg 1980
cccaggggcca tgggaaggac cttaggagtt cgatgagaga gaccatgagg ccactgggct 2040
ttcccccctc caggctctct ggggtgtcac ccttaacttt aactcttggg cctccaataa 2100
gtgtcccata ggtgtcttgg caggcccacc tgcctgggat gtggctctg tgtgtgctg 2160
tgtgggcaca ggtgtgagtg tgtgagtga agttaccaca tttcagtcct tctctgctgc 2220
aactaagtc gcaacgcc

```

<210> 34
 <211> 536
 <112> DNA
 <113> Homo sapiens

<120>
 <121> misc_feature
 <123> Incyte ID No: 2616327CB1

```

<400> 34
gcatgaactt ggggggtcagc atgctgagga tctcttctct cctggatgta ggaggagctc 60
aagtgtctgg aacaggcaag acccctgggg ctgaaattga ttccaagta cccctcctcg 120
ggacgtctgt ggggtgtctg atatctgctg gcttcttggc cctgaagatc tgcattgaca 180
ggaggcactt atttgagagc gactcttccg acctgaaaag cagccttggg ggctcagtg 240
acaccatccc gctaaagaag agagcccaa ggcgaaacca caatttctcc aaaagagatg 300
cacaggctgt tgagctgtag gtgagcagtg acgtgaagag gggttctagc cccgtggaaa 360
acagcccatg gttaacatct caggatgttc tgcattcaaa caccaagggc tggtaatgaa 420
ctttcacatg gactgaatat tggaggcaaa taatagaagg aatagaatat acagtgcctc 480
tgtcttgaag gaaaatatca tgcctcttct ggaagaaacg gactgcacag aggaag 536

```

<210> 35
 <211> 2177
 <112> DNA
 <113> Homo sapiens

<120>
 <121> misc_feature
 <123> Incyte ID No: 2991370CB1

```

<400> 35
cgggaggctc gaggccagcc cgggaccggg gctgggagca agcaggcggc ggcgcggggc 60
gcagaggcgg cagcgagcgc ccgcttccca cgcctcctagg cggcgggggc gagagcgggg 120
ggatggctcc gagcgctgac ccggcatgt ccaggatgtt accgttctct ctgctgctct 180
ggtttctgcc catcactgag ggtcccagc gggctgaacc catgttctct gcagtcacca 240
actcagttct gctcctgac tatgacagta atccccacca gctcaactat ggtgtggcag 300
ttactgatgt ggaccatgat ggggactttg agatcgtcgt ggcgggggtac aatggaccca 360
acctggttct gaagtatgac cgggcccaga agcggctggg gaacatcgcg gtcgatgagc 420
gcagctcacc ctactacgcg ctgcgggacc ggcaggggaa cgcatttggg gtcacagcct 480
gcgacatcga cggggacggc cgggaggaga tctacttctt caacatcaat aatgccttct 540
cgggggtggc cagctacacc gacaagttgt tcaagttccg caataaccgg tgggaagaca 600
tcttgagcga tgaggtcaac gtggcccggt gtgtggccag cctctttgcc ggaagctctg 660
tggcctgtgt ggacagaaag ggctctggac gctactctat ctacattgcc aattacgct 720
acggtaaatgt gggccctgat gccctcattg aaatggaccc tgaggccagt gacctctccc 780
ggggcattct ggcgctcaga gatgtggctg ctgaggctgg ggtcagcaaa tatacagggg 840
ggcagggcgt cagcgtgggc cccatctctc gcagcagtg ctcggatctc ttctgcgaca 900
atgagaatgg gcttaacttc cttttccaca accggggcga tggcaccttt gtggacgctg 960
cggccagtg cgggtgtggac gacccccacc agcatggggc aggtgtcgcc ctggctgact 1020
tcaaccgtga tggcaaagt gacatcgtct atggcaactg gaatggcccc caccgctct 1080
atctgcaaat gagcaccat gggaaggctc gcttccggga catcgctc cccaagttct 1140
ccatgccctc cctgtccgc acggctcatc ccgcgcact tgacaatgac caggagctgg 1200
agatctctct caacaacatt gctaacgca gctcctcagc caaccgctc ttccgctca 1260
tccgtagaga gcacgggagc cccctcctcg aggagctcaa tcccggcgac gccttggagc 1320

```



cagaggggacg	gggcacacagg	gggtgtgtgtga	cggatttoga	cggadadggg	atgctgggac	1380
ccatcttgttc	ccatggagag	ccatgtgtgtc	apcgcgtgtc	cgtcttccgg	ggcaatccagg	1440
gcttcaaaac	caactgggtg	cgagtgtgtg	caggaccccg	glttgggggg	tttggccagg	1500
dagctaaaggt	cgtgctctac	accaagaaaga	gtgggggccc	cctgaggatc	atcgacgggg	1560
gctcaggtta	cctgtgtgag	atggagcccg	tggacacact	tgggttgggg	aaggatgaag	1620
ccagcagttt	ggaggtgaag	tggccagatg	gcaagatggg	gagctgggaat	gtggccaggc	1680
gggagatgaa	ctcagtggtg	gagatctctc	acccccggga	tdaggacaca	cttcaggac	1740
cagcccccact	ggagtgtggg	caaggattct	cccagcagga	aaatggccat	tgcattggac	1800
ccaatgaatg	catccagttc	ccattcgtgt	gccttggaga	caagcccgta	tgtgtcaaca	1860
cctatgggaag	ctacaggtgc	cggacccaac	agaagtgcag	tgggggctac	gagccccaac	1920
aggatgggac	agcctgcgtg	ggctgggtgg	gccttgtgtt	gaagatagtg	acaccacaag	1980
ttggggaagag	ccttgggtccc	tgaatcactg	aatcactggc	ttgaatcacc	gcctgggaata	2040
cctgttgtatc	aggaacacctt	acctggaaact	tcattgagca	ggatacaaaa	ttctatttga	2100
ttaagctatt	aatacattaa	gatttggggg	tgtacactta	cataataaat	ccccatttcc	2160
tcttgaaaaa	aaaaaaaa					2220

<210> 36
 <211> 2043
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3031062CB1

<400> 36						
cgccacgacg	cagcgggggaa	tctgcagtag	gtctgcgggc	gatgggagtgg	tgggttagct	60
cgcgccttcg	gctctggttg	ctgttgtttc	tcctgccttc	agcgcagggc	cggcagaagg	120
agtcagggttc	aaaatggaaa	gtattttattg	accaaaattaa	caggtcttttg	gagaattacg	180
aacctatgttc	aagtcaaaac	tcagctctgt	accatgggtgt	catagaagag	gacttaactc	240
ctttccgagg	aggcatctcc	aggaagatga	tggcagagggt	agtcagacgg	aagtcaggga	300
cccactatca	gactactaag	aacagactgt	accggggaaaa	tgactgcatg	ttccccctcaa	360
ggtgttagtgg	tgttgagcac	tttatttttgg	aagtgtatcgg	gcgtctccct	gacatggaga	420
tggtgatcaa	tgtacgagat	tatctctcagg	ttcttaaatg	gatggagcct	gccatcccag	480
tcttctcctt	cagtaagaca	tcagagtacc	atgatatacat	gtatctctgt	tggacatttt	540
gggaaggggg	acctgctgtt	tggccaattt	atcttcacagg	tcttggacgg	tgggacctct	600
tcagagaaga	tctggtaagg	tcagcagcac	agtggccatg	gaaaaagaaa	aactctacag	660
catattttccg	aggatcaagg	acaagtccag	aacgagatcc	tcctattctt	ctgtctcgga	720
aaaaccccaa	acttgtttgat	gcagaatata	ccaaaaacca	ggcctggaaa	tcctatgaaag	780
atacctttagg	aaagccagct	gctaaggatg	tcctatcttgt	ggatcactgc	aaatacaagt	840
atctgttttaa	ttttcgaggc	gtactgcaag	tttccggttt	aaacacctct	tcctgtgttg	900
ccattattctt	aatgagaaag	agaacatact	agtatggaaa	tttcttaagg	gcaggaaagtc	960
atgtattttac	agatgtttat	tgagtacctg	ttatatataca	ggactaagct	gctgggatgt	1020
actagaaaaat	acgatatttg	tcactctgct	tcctatggaaa	ttttagacta	gcataatacc	1080
tggaacatag	tagatgaaaa	aaaatacatg	tggtttttagt	ggttgaaaatt	taagcatttg	1140
caagttatta	tacttatgac	tcacaaattt	acctttccac	gaattaagcc	aaaaagactt	1200
tatccagaaa	atagttgagt	agagatgaat	acataggatc	ctgcatgatt	taactttctc	1260
tttgcatttg	ggataatata	ctcaacacat	ttctctacca	catttaaaat	ttttattttat	1320
tacctgcctt	cttcccaaaa	ggattttaaag	cagctttacaa	aaatgtatag	cagggcgggga	1380
taaaataaag	aaacagacta	tacaacatgc	atactgtgaa	gtctcttcca	cttccaggta	1440
ccatcatatg	gaacctttat	cctcccacca	caaaatgtat	aattctgctt	atatctgtac	1500
ctgtgatctc	tcttcttttca	aaaacgtaga	cccttcttct	tatgaaaaaa	acaatcccc	1560
cacatatatt	ctggatggca	tcctcttcca	ctttgtctgg	gactttgggt	ctgcccactg	1620
tcctcatcat	ccttgtatcc	aggaccacag	cccttgtcat	gcagacatat	tcagcttctc	1680
cccatctccc	caacctccag	caaaagaaaa	cccttcattt	tcacattccc	cttcaggcca	1740
gattgaagat	gaagatgagg	atgagttcaa	ggatgaagac	caggatgagg	acaaggatga	1800
ggatggagtc	tagagcctcc	cagagcctgg	agaggaggcc	tcgggtcagcc	actccgttga	1860
cgtggggccac	ggtgaccac	catgaagtcc	ccactagcca	ctcgattccc	tgtctctgtca	1920
gagttgctgc	acatcacacc	agccctgtcc	aagagcagga	gtcaccacag	gctgaatgcc	1980
cacgaggagc	tctgctgaga	ctctcaaggg	agccagtga	agaaatagaa	ataaagcctg	2040
tgc						2043

<210> 37
 <211> 1743
 <212> DNA



<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3101617CB1

<400> 37

```

cagcagggtca cagccccctcg aggcgacacgc ggcgccgcgcg caccagagca gtggtacagg 60
catgggatggg aagaaatgca gcgtatggat gtccctacct cttgtattta ctttgtttac 120
ttcaggtgga ttgtggatag tatacttcat agctgtggaa gatgacaaaa ttttaccatt 180
aaattcagct gaaaggaaac ctgggtgtgaa gcctgcacca tatataagca ttgcagggtga 240
tgatcctcct gcaagctgtg tgtttagtca agttatgaac atggcagcct tctagccct 300
tgtggtagct gttctgcct tcatacaact gaaacccaag gttttaaacc cgtggctgaa 360
tattagtggg ttgggtggct tgtgtctggc ttccctcgga atgaccttac ttggttaatt 420
tcagctcaca aatgatgaag aaatccataa cgtcggaact tccttgacct ttggatttgg 480
cacattgacc tgctggatcc aggcctgcct gacactcaag gtcaacatca agaataaagg 540
acggagagtt ggaattccac ggggttattct gtccggcatct atcactctct gtgtggctct 600
ctacttcata ctcatggccc aaagcatcca catgtatgca gccagggtcc agtggggcct 660
ggcatgtgtg tccctgtctt attttggcac ctttgcctgt gagttccggc attaccgcta 720
tgagattgtt tgctctgagt accaggagaa ttccctaagc ttctcagaaa gcctgtcaga 780
agcttctgaa taccagactg accagggtga aaccatcagt ttttccttgc ttgtgagggtg 840
ggtgtgacag tgggggaggg gccagtagga cacaactaca ggacttgaca tagaacctca 900
tttcacacac acacacacac acacattcat ggccacattt gccaaatgag cttttcaggg 960
cgagttattt ctttaatgaa aaagcacaag ccttatgtgt tcgaaatata cgtgtttaca 1020
ctgaaaatat atgcacgaca gagcaagaag cttgtgcatg atcacttctt atccgtcccc 1080
ttcccagcac tccctcctct tcccattctc tcccatgttc tcaagcacc caccagtag 1140
ggcaggccaa atgttccctg ggagtaatgc caactccga cgttgccctc aggtccaaa 1200
ggcttggaac cagctcgtga ggaagtctg aatctggcac taatattctt gagtggataa 1260
tagtgtatca tagaatagga cggaaattgt attgagatgt gacctgtgt cgcctgtgga 1320
aaggcatagt gagaagaact tcccacgaa agcccccttc atcgttgttc agtggctggc 1380
tgtgtggatc ccaggagaga catatgccac agactgtgag agcctgtgta 1440
tctggacttg atgcactgtg actgagaatg atttccaaat gtgaatatgt gtagggaact 1500
ggtctatcag gcctggaaca agatgggggc agtgaaggta tggtttagtg tttgctttca 1560
tagtatgcca tgtacaatgt tttatatctt atagtttctt ttaagtaact accatgagtc 1620
tctctaagcc tcatggacaa agatgtagac caaatgcaag agctgagctt gctttgggtt 1680
caaccatgat caaagaaaaa ctgaggtcac ctgcaggctt acgtgggaag ctaagacaat 1740
atc

```

<210> 38

<211> 1306

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3216178CB1

<400> 38

```

ctgcaaaagtt cctgtgagcg ctgtcatttt gtcactctgg tttttcagat ttttccctg 60
gaggctggag ttccaggat gtcaaaatta cctctgcttg ggtgagctat ttcaagcagc 120
tgggatacct gtgtcactcc tgctgtctgc cagtactgc ccagggtgtct gctgggttct 180
ccccaggagt agggaggaac cagggtgggt ggctgggatg ggtggatatt taaagaccag 240
gccttggaag ctgcagcact tctatctctg cttgatgctt gctgccacgt ggtgtgtcct 300
cctcctcctg ctgtggctga gccttgggtg gaagacaggg agctgtctcc aaocccagaa 360
cctttgctgt cttgggacgg atcaccactg caagagggga agttgctact gtgatgaatt 420
ctgccatgtg gcaccagact gccaccaga ccacagtgtc ctctgcaacc ctgcttctca 480
gatgaccaag atggtgctgc agatggtgct gaggatggag aaccacacca gcccgcctag 540
gagccacctg gactggatgc agagcatggt gagctccctg caggttctctt gagaaggggt 600
ggatggcagc ctgtccttg cctttgtgct cctcaggccc caaagtcagg gaaccagaa 660
aagaaagggg cctgactag ggcagagctc cactgcaatg attgttttag ggttaggagc 720
caggattgcc gtctgtggac actgaaattt gaatctcata tacttttgtg acaaaacatt 780
cttccctctt tgttctctc ctaccatcta aaaatgtaga aaacattctt agcctatgag 840
ttgcacaaaa acaggcagtg gccagatttg gcccatagac catagtttgc tgacttctgc 900
cctaaatcat cctccatttc tttcctctg ttgtcctgtt actgacaaag ccactttccc 960
taaaatgggg tctttccctg tttgggtcca tgaagccaat atgcaaaacc gaaagtgagc 1020
ctcaagcagt gcaggcttta tttgggtggc atggaattga gaagtgagag cttggctcac 1080

```



```

aaatcaactt ttctgctcgt gagaccacagg aagtcacaga tacaggggcat ctttagtgaa 1140
ggggctgagg attaaaagca aggggaggag tggccagggt caatgggcca cccccataaa 1200
cccagaactt tgggaggcca aaatgagagg attgctgaga ccaggagtgc gagaccatcc 1260
tgggcaacat agtgatacac ccccatctct acaaaaaata aaatga 1300

```

<210> 39
 <211> 951
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3406803CB1

```

<400> 39
gggctggcca cactgcaggg gctgcaggaa gcaaaggatg aaactgatct tttcactaac 60
cagcataggg cacaggccaa agaaaaacttt gggtactctc ttgtgagcca gttgaagtta 120
gacttggtct ctttggaata ggtcttcctc tgcagaaata aaaacacttg tctgaaagag 180
gctgtaacta tctgtgggct gttgggcaag gacacttcag atactggctt tgagctcact 240
gggtcctggc cctcatggc tctgtgatca tctctgctcc acattgcagg ccattgctccc 300
tggttggtgt cagcccagga gccctccttg ggtcctggcc agactcctcc accctcgttg 360
gctgcccga accagttctag tgcttttct cccatggggg tccctggaat ctccacacac 420
ctgcccctac agagcctgca gtccagggtg ggagttaacc ctttccactt tcccagagag 480
ggagaccttg agtgaggggg aggtcaggaa acgcgggggt ggcagcatgg tgggaggagg 540
tgaaggtacg atgaccagag ccttggtgtg gcggtggtg acaaaagctga ggttatatg 600
ggtgctgtgt gcaggggcgt cctgagaaga cactcctga tggagtgtg ggggaagtgc 660
aaggcttgga ggccaatgag ggggctgatg gcacccatcc agtggagagg ctgtatttgg 720
agggcctcct gagtgggagc aaggctatgt ggatgatgag atgaggtaag tgagatggaa 780
gtggctgtac gaacagccat tgaggctgag gcaggaggat tgcctaagtc catagcccag 840
gtgtcgagac t 851

```

<210> 40
 <211> 2204
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3468066CB1

```

<400> 40
gagcagtcct tgctgggtccc gcccccgctc ggctcgcgcg caggggagcg tagtgggtcc 60
agtctccttg gcaaccactt gctcctcccc ctccgcccct ttaaccttta gggcgcgagg 120
gtgcagtata tctcgcgctc tctccccctt cccccccccc ttttcccacc ccggggcgctc 180
agggttggtct ggaccggaag cgaagatggc gacttctggc ggggctcggg cggagctggg 240
gatcgggttg tgcatattcg gctcttact actggctatt ttggcattct gctggatata 300
tgttcgtaaa taccaaagtc ggcgggaaag tgaagtgtc tccaccataa cagcaatttt 360
ttctctagca attgcactta tcacatcagc acttctacca gtggatata ttttggtttc 420
ttacatgaaa aatcaaaatg gtacatttaa ggactgggct aatgctaatt tcagcagaca 480
gattgaggac actgtattat acggttacta tactttatat tctgttatat tgttctgtgt 540
gttcttcttg atcccttttg tctacttcta ttatgaagaa aaggatgatg atgatactag 600
taaatgtact caaattaaaa cggcaactcaa gtatactttg ggatttgttg tgatttgtgc 660
actgcttctt ttagttgggtg cctttgttcc attgaatgtt cccaataaca aaaattctac 720
agagtgggaa aaagtgaagt cctattttga agaacttga agtagtcatg gtttagctgc 780
attgtcattt tctatcagtt ctctgacctt gattggaatg ttggcagcta taacttacac 840
agcctatggc atgtctgcgt tacctttaaa tctgataaaa ggcactagaa gcgctgctta 900
tgaacgtttg gaaaacactg aagacattga agaagtagaa caacacatte aaacgattaa 960
atcaaaaagc aaagatggtc gaccttggc agcaagggat aaacgcgctt taaaacaatt 1020
tgaagaaaagg ttacgaacac ttaagaagag agagaggcat ttagaattca ttgaaaacag 1080
ctgggtggaca aaattttgtg gcgctctgcg tccctgaag atcgtctggg gaattttttt 1140
catcttagtt gcattgctgt ttgtaatttc tctcttcttg tcaaatttag ataaagctct 1200
tcattcagct ggaatagatt ctggtttcat aatttttggg gctaacctga gtaatccact 1260
gaatatgctt ttgcctttac tacaaacagt ttccctctt gattatattc ttataacaat 1320
tattattatg tacttttatt ttacttcaat ggcaggaatt cgaaatattg gcataagggt 1380
cttttggatt agattatata aaatcagaag aggtagaacc aggccccaag cactcctttt 1440

```



tctctgcatg	atacttctgc	ttattgtcct	tcacactagc	tacatgattt	atagtcttgc	1500
ttcccaatat	gttatgtatg	gaagccaaaa	ttacttaata	gagactaata	taacttctga	1550
taatcataaa	ggcaattcaa	ccctttctgt	gocaaagaga	tgtgatgcag	aagctcctga	1600
agatcagtg	actgttacc	ggacatacct	atctcttcac	aagttctggt	tcttcagtg	1650
tgcttactat	tttggttaact	gggcctttct	tgggggtatt	ttgattggat	taattgtatc	1700
ctgttgtaaa	gggaagaaat	cgggttattga	aggagtagat	gaagattcag	acataagtga	1750
tgatgagccc	tctgtctatt	ctgcttgaca	gccttctgtc	ttaaagggtt	tataatgctg	1800
actgaatatc	tgntatgcat	ttttaaagta	ttaaactaac	attaggattt	gctaactago	1850
tttcatcaaa	aatgggagca	tggctataag	acaactatat	tttattatat	gtttcttgaa	1900
gtaacattgt	atcatagatt	aacattttaa	attaccataa	tcatgctatg	taaatataag	1950
actactggct	ttgtgagggg	atgtttgtgc	aaaatttttt	cctctaattg	ataatagtg	2000
taaattgatt	aaaaatcttc	cagaattaat	attccctttt	gtcacttttt	gaaaacataa	2050
taaatcatct	gtatctgtgc	cttaggttct	ccaaaaaaaa	aaaa		2100

<210> 41
 <211> 570
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3592862CB1

<400> 41	
ggcgaggagc	tcggggagtc
gggtctcggc	cacggcgtgg
aaacaagttc	ttctcaagga
agttaccaga	aggctgttt
aaaaatttcc	agatgatgtt
aacctaccgc	atatggagat
attctttaac	ttcaatatgt
atcttctttc	tttatatcct
tatatatata	tatatatata
attaggaaaag	ggagaccagc
ggcgccatga	ccccatcgag
agagcggcaa	gatcacccct
ggaaagtctg	aacttgtcaa
gatgcaccag	aggattccca
aatccagtg	ccaaagaaaa
tgggaacgaa	aaggacgctg
ttttctgaat	atgtacatct
ttatgtcgtg	tagtttgtgt
tatatatata	ggggcttagg
ctgaccaata	aagaaaatat
ttgcttagct	60
tctgtgtcat	120
tctctgagga	180
acagtcctct	240
aagaagccga	300
gaaccactgg	360
aggtggaccc	420
aggggccag	480
taagtgcgat	540
atttctgatt	570

<210> 42
 <211> 802
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3669422CB1

<400> 42	
caggggtcaag	gtgaagctgg
tgggggtcac	actggctctc
cttccccagg	gatgtctctg
ccacatcccc	ctccccagac
gggaggggcct	gaagcagggc
ccccctcttc	ccgtcccccc
caatcccttc	acactctctc
aatgtgggca	ttaatTTTTT
gctgtgttca	tacctaaatt
aagggggaga	caaaaactcct
ctgcccccat	tccttcaaga
aaaaagggtta	gaagggggcc
atgatggaag	atggcattag
actggagatg	cagagtcctc
tggtgtctcg	agggcgggtga
tctagtccca	tgctgtctgc
tggagctgcc	ttttgttctt
cccagtcagg	tgagcacact
cagtggagga	aaactggccc
agccccctct	cccttcccc
ccccatcccc	ccaagataca
gactgcagct	ctgcttctcc
ttctggaagg	ggacagtga
actctcaacc	tcacaccaac
ggagaccctt	tggggacaag
cctctctgaa	ggcccccaact
ctacaccacc	ctgcgctttc
ag	ccgagatgaa
ggagcctggg	60
tgccctctcc	120
agccccacga	180
agccctcaga	240
cacccccaca	300
cttcttcaac	360
ctcttctgctg	420
tgggtggcaa	480
acaggaggga	540
atcactctct	600
tgtaggaat	660
ctacaatcca	720
cataccacga	780
	802

<210> 43
 <211> 693
 <212> DNA



<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3688740CB1

<400> 43

gttggttttaa	tgggattgtg	gaagagaatg	actccaatat	ttggaagtgc	tggtacacca	60
atcagccacc	gtccaagaac	tgcacacatg	cttacctgtc	tcctgacccc	ttcatgagag	120
gagagcacia	ctcgacctcc	tatgactctg	cagttattta	ccgtgggttc	tgggcagtc	180
tgatgtccct	gggggtagtt	gctgtagtca	tgcgaagctt	tttgatcctc	tgtgcagccc	240
ccctcgccag	ccattttctc	tacaaagctg	ggggaggctc	atatattgct	gcagatggaa	300
ttctctctct	ctgttactca	agcctctcaa	agtccttatt	gtcccagcct	ctgcgtgaaa	360
cgtcttcagc	catcaatgac	atctcactcc	ttcaagccct	tatgccactg	ctgggatgga	420
ccagtcaactg	gacctgcatt	acagtgggct	tatatggacc	tttacctcta	cacttgatata	480
taactcgttt	tcccatcttag	ttgcaagaca	cttgggaagca	cagaccaagg	cttacatttg	540
tgttttaattg	ttttctctgt	aaatgcttta	tgcctaaatg	ttctgtact	actcttcttt	600
ccaaatcctt	attgttttaa	agtctctctc	ctactatacc	atgccttata	aatattgatt	660
gaatgaatgg	atgaaatgca	tacctgctta	tag			693

<210> 44

<211> 1212

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3742589CB1

<400> 44

ccctcgaggc	aacttgcctt	tctcaaacat	ggcgcgccag	gcgcctctgg	aagggaaccg	60
ctctggggcc	cgcctttgat	ctcgttggtg	gggctggggg	atgagagctg	caccgcgcgg	120
gacaagtccg	cggcgggcgc	cgacggagca	gaacagagag	catggagctg	gagaggatcg	180
tcagtgcagc	cctccttgcc	tttgtccaga	cacacctccc	ggaggccgac	ctcagtggct	240
tggatgaggt	catcttctcc	tatgtgcttg	gggtcctgga	ggacctgggc	ccctcggggc	300
catcagagga	gaacttcgat	atggaggctt	tcaactgagat	gatggaggcc	tatgtgcctg	360
gcttcgcccc	catccccagg	ggcacaatag	gggacatgat	gcagaagctc	tcagggcagc	420
tgagcgatgc	caggaacaaa	gagaacctgc	aaccgcagag	ctctggtgtc	caaggtcagg	480
tgcccactct	cccagagccc	ctgcagcggc	ccgaaatgct	caaagaagag	actaggtctt	540
cggctgctgc	tgctgcagac	acccaagatg	aggcaactgg	cgtgaggag	gagcttctgc	600
caggggtgga	tgtactcctg	gaggtgttcc	ctacctgttc	ggtggagcag	gcccagtggg	660
tgctggccaa	agctcggggg	gacttggaag	aagctgtgca	gatgctggtg	gagggaaagg	720
aagagggggc	tgcagcctgg	gagggcccca	accaggacct	gcccagacgc	ctcagaggcc	780
cccaaaagga	tgagctgaag	tccttcctcc	tgcagaagta	catgatgggtg	gatagcgcag	840
aggatcagaa	gattcaccgg	cccatggctc	ccaaggaggc	ccccaaagaag	ctgatccgat	900
acatcgacaa	ccaggtagtg	agcaccaaaag	gggagcgatt	caaagatgtg	cggaaacctg	960
aggccgagga	gatgaaggcc	acatacatca	acctcaagcc	agccagaaaag	taccgcttcc	1020
attgaggcac	tgcgcggact	ctgcccagac	cttctaggct	cagatcccag	agggatgcag	1080
gagccctata	cccctacaca	ggggccccct	aactcctgtc	ccccttctct	actcctttgc	1140
tccatagtgt	taacctactc	tcggaagctgc	ctccatgggc	acagtaaagg	tggcccaagg	1200
aaaaaaaaaa	aa					1212

